Long-Term Risks of Risk-Based Allocation

Jivat Neet Kaur*
University of California, Berkeley
jnkaur@berkeley.edu

Jane H. Lee*
Yale University
jane.h.lee@yale.edu

Manolis Zampetakis*
Yale University
manolis.zampetakis@yale.edu

Abstract

Algorithmic risk predictors are increasingly being used to inform resource allocation in healthcare and other high-stakes domains. Common in many such settings, the public decision maker delegates service provision to private organizations and compensates them according to each individual's predicted cost of service. We show that this common allocation mechanism can fail when organizations engage in *favorable selection*, i.e., strategically choosing individuals whose actual costs are below the payment model predictions. Motivated by health insurance payments in the Medicare Advantage (MA) program, we provide a theoretical framework to study the interaction between a decision maker and a strategically selecting service provider. Our theoretical framework predicts two empirical patterns observed in Medicare Advantage that have become central concerns: enrollment in the private program expands, while government payments rise above the counterfactual cost of direct public provision. Guided by our theoretical framework, we then propose a reformed payment policy that is implementable under the government's data-access constraints. We prove that its equilibria eliminate overpayment while preserving profits for efficient organizations. We validate the framework in synthetic and semi-synthetic experiments that illustrate the advantages of our reformed policy, over the status quo.

1 Introduction

Algorithmic risk predictions are increasingly used to allocate resources in high-stakes settings, from deciding which patients receive expanded health care [1], to which households are eligible for targeted social benefits [2, 3], to which jobseekers receive employment assistance [4]. A common pattern across many of these settings is that the public decision maker does not deliver the service itself: it contracts with private organizations to do so, and pays each organization an amount calibrated to the expected cost of serving each individual. The motivation for this delegation is that private organizations are often expected to deliver services more efficiently, with greater specialization, or with broader reach than direct public provision; the motivation for paying at predicted cost is to make the arrangement budget-neutral relative to direct provision in expectation, while still leaving the organization with an incentive to deliver services efficiently.

Our main motivation comes from the *Medicare Advantage* (MA) program, an alternative private-plan option within Medicare, the U.S. federal health insurance program primarily for people aged 65 and older. Under *traditional Medicare*, also known as *fee-for-service* (FFS), the government directly reimburses medical facilities for each service they deliver to a beneficiary.

*Authors are listed alphabetically

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

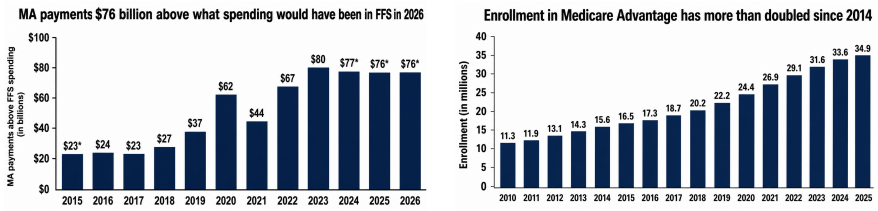

Figure 1: The left panel, reproduced from a MedPAC status report to Congress, shows payments to MA plans have consistently exceeded what spending would have been under traditional fee-for-service (FFS) Medicare. The right panel shows that Medicare Advantage (MA) enrollment has grown steadily over time. (See [10].)

Under MA, the government instead contracts with a private insurance plan: in exchange for a fixed annual payment per enrolled beneficiary, the plan takes on responsibility for delivering that beneficiary's covered care for the year.

The central design question for the government is how much to pay each plan per enrollee. Healthcare costs vary widely across individuals, e.g., sicker beneficiaries can cost an order of magnitude more to treat than healthier ones, so paying a uniform amount per enrollee would create a strong incentive for plans to enroll only the healthiest. To prevent this, the government uses a *risk-adjustment model*: a regression that maps each beneficiary's observable characteristics (demographics, recorded diagnoses, etc.) to a predicted annual cost, and the plan's payment for each enrollee is based on that prediction. The model currently in use, called the CMS-HCC model, uses health-condition variables known as *hierarchical condition categories* (HCCs), and is fit on the FFS population, which is the only population whose individual-level realized costs the government directly observes [5, 6, 7].

The same structure recurs elsewhere in U.S. health policy. Medicaid, a separate federal–state program covering low-income individuals, has a parallel program called *Medicaid managed care*, which similarly pays private plans fixed per-enrollee rates [8]. The same broad logic also underlies the risk-adjustment system used in the health insurance marketplaces created by the Affordable Care Act (ACA) [9]. Two empirical patterns in the MA program are by now well documented.

(Pattern 1) Per-enrollee government payments rise above counterfactual costs. MedPAC estimated that in 2025 Medicare would pay MA plans roughly $84 billion (20%) more than it would have spent on comparable beneficiaries under FFS, and attributes on the order of $44 billion of this gap to *favorable selection*, i.e., to the systematic enrollment of individuals whose counterfactual FFS costs are below what the payment model predicts [11, 12, 13] (see Figure 1 (left)). This pattern has persisted across successive refinements of the risk-adjustment model [14, 15, 16, 17, 18], and is further amplified when plans more intensively document diagnoses for selected enrollees [16, 19, 18].

(Pattern 2) Enrollment grows. The share of Medicare beneficiaries enrolled in MA rose from roughly one-quarter in 2010 to over half by 2024, and is projected to keep expanding [20, 8, 13] (see Figure 1 (right)). This growth exacerbates the consequences of Pattern 1 and has also raised broader concerns about the current payment system [12, 13].

Despite the effort to get a better handle of Patterns 1 and 2, there is currently no theoretical explanation of these patterns that can be derived from a model of the interaction between the decision maker and the strategically selecting providers. In the meantime, substantial overpayments persist despite successive revisions to the risk-adjustment model and broader efforts to improve the accuracy of risk-adjusted payments [21, 13]. This motivates the central question of this paper:

Is there a tractable theoretical framework that: (i) *predicts* Patterns 1 *and* 2*, and* (ii) *enables the principled design of an alternative policy that has provably better properties?*

In this manuscript, we present such a theoretical framework. We provide analyses that explain these patterns and then propose a principled alternative policy to normalize them.

1.1 Our Contributions

In this work, we study a setting where a decision maker introduces a new program alongside an existing policy, contracting with organizations to provide services and setting payments to reflect the

expected costs. The decision maker aims to encourage efficient service provision, while organizations may have incentives to preferentially serve individuals whose costs are overestimated by the payment function. Formally, the decision maker uses data $(x, y) \sim \mathcal{D}$, where x denotes an individual's features and y the cost of serving that individual under the existing program, to estimate a payment or allocation function $f_\theta(x) \approx \mathbb{E}[Y \mid X = x]$. The organizations may incur a different cost $y' \in \mathbb{R}$ when serving that member in the new program, reflecting differences in how services are delivered. The decision maker aims for organizations to serve individuals for whom $y' < y$, since for these individuals the new program provides services at lower cost than the existing policy.

However, organizations are incentivized to serve individuals for whom $y' < f_\theta(x)$. Organizations then act strategically to affect participation in such a program through *favorable selection*. We formalize this as a selection rule $\pi : \mathcal{X} \times \mathbb{R} \to \{0, 1\}$, where $\pi(x, y') = 1$ indicates that an individual with features x is served by the organization, and $\pi(x, y') = 0$ otherwise. Favorable selection may be exercised through targeted advertisement and marketing [17], differential service quality, and service design targeted at profitable individuals while discouraging unprofitable ones [22, 23].

Status quo. In practical applications, the decision maker continues to estimate payment benchmarks using observed costs from the population that remains under the existing policy. This assumes that costs in the remaining population provide a reasonable benchmark for the members served under the new program. It is further reinforced by institutional inertia and reliance on established pipelines built around those data. As a result, at any point in time, the decision maker often only observes costs y for individuals who remain under the existing policy, and therefore estimates the payment rule by:

$$\min_\theta \mathbb{E}\left[(y - f_\theta(x))^2 \mid \pi(x, y') = 0\right]. \tag{1}$$

Policy reform. We propose a reformed policy where the decision maker estimates the payment rule using costs y for the population selected into the new program. We note that this is a counterfactual (and hence unobserved) quantity and not directly accessible to the decision maker. Nonetheless, we will show that we can estimate this quantity given distributional assumptions on (x, y) (Section 4.3).

$$\min_\theta \mathbb{E}\left[(y - f_\theta(x))^2 \mid \pi(x, y') = 1\right]. \tag{2}$$

Theoretical analysis. Our main theoretical contributions are (1) providing a framework to formalize the qualitative goals and empirical observations in risk-based allocation (Section 2), (2) the resulting theoretical insights of status quo policies through analyzing such framework (Section 3), and (3) proposing an implementable alternative policy which mitigates undesirable behavior (Section 4) by leveraging recent results in truncated regression.

Empirical results. We simulate the dynamics we formulate on synthetic and semi-synthetic environments to validate the qualitative interpretations of our theoretical results. In the semi-synthetic setting, both the covariates x and the costs y are constructed to reflect realistic features of the Medicare population, including co-occurring health conditions and heavy-tailed healthcare spending. Although our theoretical assumptions do no hold in this environment, we nevertheless observe qualitatively similar patterns to those predicted in the idealized setting (Section 5).

Insights and discussion. Through our theoretical analysis and empirical evaluation, we show that the status quo policy in risk-based resource allocation programs can lead to undesirable long-term consequences. More broadly, our analysis illustrates that improving predictive accuracy alone need not resolve these concerns as long as the decision maker's optimization problem remains unchanged. This highlights a disconnect between existing approaches to risk adjustment and the underlying design problem of aligning incentives with the planner's social welfare goals. Our results therefore suggest policy redesign, rather than prediction alone, as the central lever for improvement.

1.2 Related Work

Economics and policy. Our work relates to a large body of work on selection in insurance markets. Early work on adverse selection shows how individuals can self-select into contracts based on private information about their own risk, leading to distorted pricing, inefficient sorting, or market unraveling [24, 25, 26]. Our setting differs in that organizations strategically choose which individuals

to attract or serve in response to a learned payment rule. Risk adjustment in health insurance is introduced with the goal of alleviating such selection by compensating plans appropriately for the predictable cost differences in individuals they serve. However, due to private information about their own efficiencies and additional information about enrollees, private plans can continue to favorably select individuals [14, 15, 17, 12]. In a setting similar to ours, [27] study the role of big data in reducing selection by private providers in government-run capitation schemes, such as Medicare Advantage. In contrast, we study long-term dynamics in this work, while focusing on existing regression pipelines and interventions available to the decision maker. Medicare Payment Advisory Commission has proposed alternative approaches to setting MA benchmarks that rely less exclusively on FFS spending, in response to concerns about a declining FFS population [12]. While those proposals rely on the availability of MA spending data, our proposed policy is implementable under existing data access constraints and still provably eliminates overpayments.

Learning in non-zero-sum games. Our framework casts the interaction between decision maker and organization as a non-zero-sum two-player game [28]. This places it in a broad problem space spanning heterogeneous preferences and incomplete information [29, 30], structures [31, 32], solution concepts [33, 34, 35, 36], and guarantees on learning and efficiency [37, 38]. The timing aspect of our setting could be compared to Stackelberg games [39], as the decision maker commits first to publishing a payment model and the organization responds. However, the leader in our framework does not respond by anticipating the follower's response and instead responds myopically. We analyze fixed points (a Nash equilibrium), which may not exist for general games [40].

Truncated regression and sample selection. Learning in the presence of truncated or censored data has a long history with a large body of literature. For a broad background on sample selection and correction methods, we refer the reader to [41, 42], and [43], as well as textbook discussions in, e.g., [44] or [45] for a perspective from the missing-data literature. [46] introduced the canonical regression-with-selection model, with a large follow-up literature (e.g., [47, 48, 49]). Outcome-censored regression is a special setting which has its own long history (e.g., [50]), which has recently received significant attention in theoretical computer science which focuses on computational efficiency [51, 52, 53, 54, 55, 56].

Applications. Examples of favorable selection extend beyond health insurance. In education, work on school vouchers and private-school choice studies cream skimming and sorting, where private providers may attract relatively advantaged or lower-cost students [57, 58, 59]. [60] study outsourced welfare-to-work services and show that performance-based contracts can induce cream-skimming and parking as private providers prioritize unemployed job seekers who are easier to place.

2 Formal Framework

We study the long-term interaction between a decision maker who specifies the payment rule and organizations that choose their selection rule. The decision maker publishes a payment rule $f_\theta : \mathcal{X} \to \mathcal{Y}$ parameterized by $\theta \in \Theta \subseteq \mathbb{R}^d$, mapping an individual's feature representation $x \in \mathcal{X} \subseteq \mathbb{R}^d$ to the predicted outcome $f_\theta(x) \in \mathbb{R}$. We will consider linear predictors $f_\theta(x) = \theta^\top x$. We implicitly include an intercept by appending 1 to each feature vector. While the linear specification is a commonly studied setting amongst theoreticians, this is not merely a modeling assumption: many models used in practice—e.g., payment models in healthcare including the risk adjustment model used to allocate payments to MA plans—are based on least squares regression [7, 6].

We start by modeling the organization's strategic behavior as the following selection problem.

Definition 1 (Organization's selection problem)**.** *The organization seeks to identify a selection rule* $\pi : \mathbb{R}^d \times \mathbb{R} \to \{0, 1\}$ *that maximizes its expected net payoff given the decision-maker's payment rule* $\theta \in \mathbb{R}^d$.

$$\max_{\pi:\, \mathbb{R}^d \times \mathbb{R} \to \{0,1\}} \mathbb{E}[\pi(x, y')(x^\top \theta - y')]. \tag{3}$$

Lemma 1. *The optimal selection policy* $\pi^\star : \mathbb{R}^d \times \mathbb{R} \to \{0, 1\}$ *to solve the optimization problem in Definition 1 is equal to* $\pi^\star(x, y') = \mathbb{1}\{y' < x^\top \theta\}$.

Remark 1. Organizations may incur costs in carrying out selection, for example through screening or targeted advertising. In our framework, we absorb such costs into y' .

Decision maker's objective. In practice, the decision maker seeks to identify a payment rule $\theta \in \mathbb{R}^d$ that serves multiple goals. For each selected individual with features x, the rule determines a predicted payment $x^\top \theta$. This payment should reflect the true costs for providing the service, while at the same time preserving incentives for organizations to provide it, for example by allowing them to retain some cost-saving profits. We formalize these objectives in our framework through the following two quantities of interest:

$$\begin{aligned} \text{Decision maker's overpayment:} \qquad & \mathbb{E}[(x^\top \theta - y) \mid \pi(x, y') = 1] \\ \text{Organization's margin:} \qquad & \mathbb{E}[(x^\top \theta - y') \mid \pi(x, y') = 1] \end{aligned}$$

Data access constraints. Often, the decision maker does not observe data for the selected population with $\pi(x, y') = 1$. For individuals served by the private organizations, the realized costs y' are often opaque to the decision maker. Moreover, the decision maker does not observe counterfactual labels y for this selected population once they participate in the program and realize costs y' instead. These constraints likely contribute to the choice of using the status quo dynamics described above.

Remark 2. The overpayment objective above aligns with the treatment of favorable selection in literature [11]. Specifically, favorable selection analyses ask whether risk-adjusted payments for the selected population exceed what Medicare would have spent on those same beneficiaries under FFS, absent any intervention by MA plans. Efficiency gains, $y' < y$, are not counted as overpayment: they generate a legitimate margin for the organization as long as payments do not exceed the decision maker's counterfactual costs. We discuss these design choices in detail in Appendix B.

3 Status Quo Policy

In practice, the decision maker only has access to (x, y) from the population such that $\pi(x, y') = 0$. Motivated by the assumption that average costs in the remaining population provide a reasonable benchmark for members served under the new program, the organization and decision maker's problems, respectively, are given by

$$\max_{\pi:\, \mathbb{R}^d \times \mathbb{R} \to \{0,1\}} \mathbb{E}[\pi(x, y')(x^\top \theta - y')] \qquad \text{and} \qquad \min_{\theta \in \mathbb{R}^d} \mathbb{E}[(1 - \pi(x, y'))\,(x^\top \theta - y)^2]\,. \tag{4}$$

3.1 Warm Up: Perfect Proxy $y = y'$

We can say more in the special case where $y = y'$. In this case, the service costs are the same regardless of whether individuals are served by the organization or by the decision maker. From the decision maker's perspective, there is therefore no inherent benefit to having individuals served by the organization. Ideally, the decision maker would hope to pay exactly y for the individuals such that $\pi^\star(x, y') = 1$. However, this is far from what happens.

Consider the following finite-sample version of the dynamics described in Equation (4). Suppose the game starts with $\{(x_i, y_i)\}_{i \in [n]}$ for a sufficiently large n. The objectives of the two players are

$$\max_{\pi} \frac{1}{n} \sum_{i \in [n]} \pi(x_i, y_i) \cdot \left(x_i^\top \theta - y_i\right) \quad \text{and} \quad \min_{\theta} \frac{1}{n} \sum_{i \in [n]} (1 - \pi(x_i, y_i)) \cdot \left(y_i - \theta^\top x_i\right)^2 .$$

The best response for each player in every round t are given by

$$\pi^t(x_i, y_i) \coloneqq \mathbb{1}\left\{x_i^\top \theta^t > y_i\right\} \qquad \text{and} \qquad \theta^{t+1} = X_{\overline{\pi^t}}^{+} y_{\overline{\pi^t}}\,,$$

where $X_{\overline{\pi^t}}$ is the design matrix taking the rows such that $\pi^t(x_i, y_i) = 0$ and $y_{\overline{\pi}}$ the corresponding labels. The pseudo-inverse $X_{\overline{\pi^t}}^{+} = \left(X_{\overline{\pi}}^\top X_{\overline{\pi}}\right)^{-1} X_{\overline{\pi}}^\top$ when $X_{\overline{\pi}}^\top X_{\overline{\pi}}$ is invertible. There are many solutions when $X_{\overline{\pi}}^\top X_{\overline{\pi}}$ is not invertible, but we will take the canonical solution $X_{\overline{\pi^t}}^{+} y_{\overline{\pi^t}}$ in this case.

Lemma 2. *Suppose the min player returns the canonical OLS solution when the design matrix no longer has full column rank. Then in every round t for which $X_{\overline{\pi^t}}$ is non-empty,*

$$\widehat{\mathbb{E}}\left[X_{\overline{\pi^t}}\left(\theta^{t+1} - \theta^t\right)\right] \coloneqq \frac{1}{n^t} \mathbf{1}^\top \left(X_{\overline{\pi^t}}\left(\theta^{t+1} - \theta^t\right)\right) \geq 0\,,$$

where $n^t = |\{(x_i, y_i) : \pi^t(x_i, y_i) = 0, i \in [n]\}|$. The inequality is tight iff when $X_{\overline{\pi^t}} \theta^t = y_{\overline{\pi^t}}$.

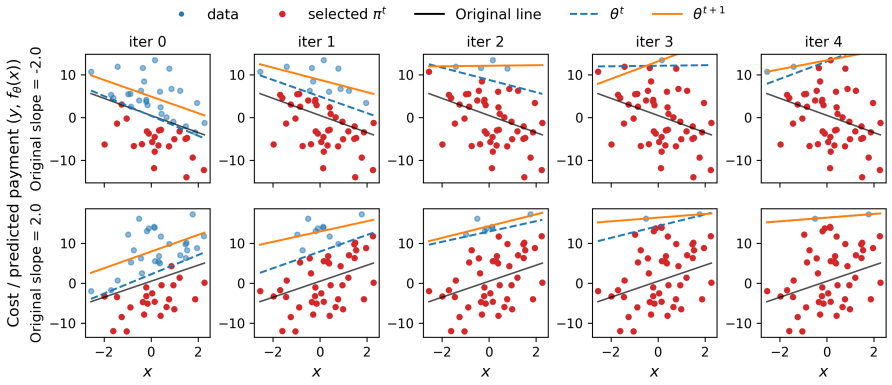

Figure 2: Illustration of the status quo dynamics when $y = y'$. At each iteration, the organization selects points below the current payment line and the decision maker refits on the remaining points, pushing predictions upward until nearly all points are selected. See Appendix E for details.

Proposition 1 (Equilibrium when $y = y'$ under (4))**.** *(Upper envelope of points lies exactly on a line.) Suppose there exists a non-empty set S such that for all $x_i \in S$, $x_i^\top \widetilde{\theta} = y_i$ exactly (where $\widetilde{\theta} = X_S^+ y_S$) and for all other $x_i \in \mathcal{X}$ for $i \in [n]$, it holds that $x_i^\top \widetilde{\theta} > y_i$. Then a fixed point is*

$$\pi^\star(x_i, y_i) = 1 - \mathbb{1}_S(x_i) \qquad \textit{and} \qquad \theta^\star = \widetilde{\theta}\,.$$

Remark 3 (Cyclical equilibria)**.** The above proves that if we reach such a state, we will remain there in every subsequent round. Due to the tie-breaking for θ, the response π^t is deterministic for each corresponding θ^t. However, it is possible for the dynamics to have a cycle length greater than 1 if the sequence $\pi^t, \pi^{t+1}, \ldots$ cycles without reaching this fixed point behavior. Figure 2 illustrates convergence to a fixed point, but convergence is not guaranteed in general. In the finite-sample setting, nontrivial cycles can be ruled out by enforcing a rule like $\pi^{t+1}(x, y) = \pi^t(x, y)\,\mathbb{1}\left\{x^\top \theta^t > y\right\}$, which prevents previously unselected individuals from becoming selected.

The proofs appear in Appendix C.3; a corresponding population result is deferred to Appendix D.1.

Phenomenon 1 (**The decision maker's payment is as large as possible while enrollment increases**)**.** The above result shows that in this special case, average predictions increase on the remaining data. At a fixed point, the payment line lies above all selected points, while the remaining unselected points lie exactly on the line. In finite samples, this results in the equilibrium line lying on the upper envelope of the given samples , while the max player picks nearly all the available data. (See Figure 2). These effects may help explain, at least partially, what is observed in practice (Figure 1).

Phenomenon 2 (**Effects of selection bias can be catastrophic**)**.** Despite the hope that available data may suffice for the decision maker's goals, we show that naively estimating costs for the unseen cohort from such data is not benign and can lead to substantial overpayments.

3.2 General case

Problem setup. Under the simplifying assumption that the true cost of providing services in the existing policy is given by $y = x^\top \beta + \beta_0 + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, covariates are Gaussian independent of noise, and relating y and y' via $y' = y - \lambda(x)$, where $\lambda(x) = x^\top \alpha + \alpha_0$ is affine, we can prove that these dynamics lead to an equilibrium wherein the decision maker will overpay on average. Without loss of generality, we can assume $x \sim \mathcal{N}(0, \Sigma)$ or even $x \sim \mathcal{N}(0, I)$ since we can center the distribution or standardize it via linear transformation. For this analysis, we write the payment rule as $f_{\theta,\theta_0}(x) = x^\top \theta + \theta_0$, where θ_0 denotes the intercept, previously absorbed into the feature vector.

The Gaussianity of covariates is not always needed, but the Gaussian distribution satisfies the assumptions for each of these stated results. We refer the reader to Appendix D.2 for a discussion of the weaker assumptions under which some of these conclusions extend beyond Gaussian covariates. We also discuss implications of non-affine $\lambda(x)$ in Appendix D.3 and D.1.

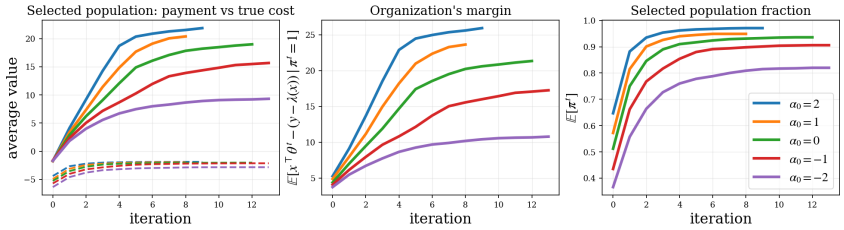

Figure 3: Dynamics of government payments, organization margins, and selected population fraction under the status quo policy. Solid and dashed lines in the left panel represent average payments and true costs, respectively. See Appendix E for details.

Theorem 1 (Status quo (see Theorem C.1))**.** *Suppose there exist* $(\theta^\star, \theta_0^\star)$ *and* $\pi^\star$ *which satisfy the fixed point equations. The following hold.*

$\pi^\star(x,y) = \mathbb{1}\left\{x^\top\theta^\star + \theta_0^\star > y - \lambda(x)\right\}$ *and* $(\theta^\star, \theta_0^\star) = \arg\min_{\theta,\theta_0} \mathbb{E}\left[(y - x^\top\theta - \theta_0)^2 \mid (x,y') \in \overline{\pi^\star}\right]$.

- ***Decision maker overpays on the selected set.*** $\mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1\right] > 0$.
- ***Organization margin on selected comes from both overpayment and advantage due to*** $\lambda(x)$**.**

$$\mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y' \mid \pi^\star] = \mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star\right] + \mathbb{E}\left[\lambda(x) \mid \pi^\star\right]$$
$$\Rightarrow \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right] < \mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y' \mid \pi^\star]\,.$$

- ***Selected population fraction.*** *Moreover, if* $|\alpha_0| < \theta_0^\star - \beta_0$, $\mathbb{E}[\pi^\star(x,y')] > 1/2$.

We illustrate these qualitative findings in Figure 3. From this, we observe the following.

Phenomenon 3 (**Organization margin is always positive, regardless of the payment model**)**.** We first remark that the organization's margin being positive on average, regardless of what $\theta^\star$ is, is natural given the definition of $\pi^\star(x,y')$. Since every point in this set satisfies $\theta_0^\star + x^\top\theta^\star - y' > 0$, it should be true of their average. This is reflected in Lemma C.1.

We also note that the organization's margin can be decomposed into the decision maker's overpayment plus the organization's cost advantage relative to direct provision. Thus we observe the following.

Phenomenon 4 (**Organization margin comes from overpayment and efficiency**)**.** There is a tension that happens between how much of the organization's margin is due to the decision maker's overpayment and the "efficiency" due to the gap $y - y' = \lambda(x)$.

This decomposition highlights why overpayment is problematic: the organization's margin can increase either because it delivers the service more efficiently, through the gap $y - y' = \lambda(x)$, or because the decision maker pays more than its own counterfactual cost. Only the first source of margin reflects cost-effective service provision. Finally, the size of the selected set depends on parameters of the fixed point in general and should not require any particular restrictions (lower or upper bounds). We emphasize that increasing enrollment alone is not an issue in itself. However, we can make some qualitative observations under special cases.

Phenomenon 5 (**Selected set is the majority, even when average advantage is 0**)**.** If α_0 is sufficiently small we see that a majority of population will be selected. In particular, since $\lambda(x) = \alpha^\top x + \alpha_0$ and $\mathbb{E}[x] = 0$, this means that if $\alpha_0 = 0$ the population average $\lambda(x)$ is roughly 0. We may want to enroll roughly half (those with $y - y' = \lambda(x) > 0$), but we select more.

Remark 4 (Existence and convergence)**.** Existence of such fixed points is detailed in Appendix C.1.3. We can make statements about local or global convergence under some additional (perhaps strong) assumptions about the trajectory of the dynamics. In full generality, such dynamics may not always converge. We defer this discussion to Appendix D.3.

4 Policy Reform

Given the behaviors we observed under the status quo policy, we propose the following alternative.

$$\max_{\pi:\,\mathbb{R}^d\times\mathbb{R}\to\{0,1\}} \mathbb{E}[\pi(x,y')(x^\top\theta - y')] \quad \text{and} \quad \min_{\theta\in\mathbb{R}^d} \mathbb{E}[\pi(x,y')\,(x^\top\theta - y)^2]\,. \tag{5}$$

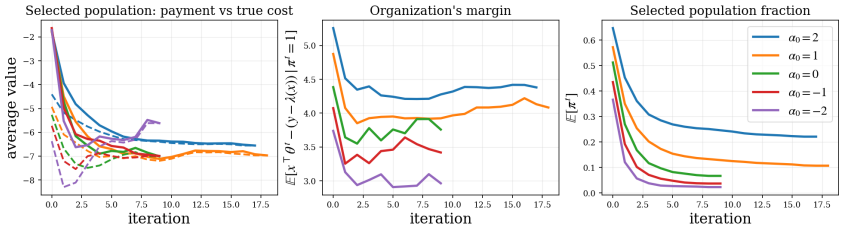

Figure 4: Dynamics of government payments, organization margins, and selected population fraction under the reformed policy. Solid and dashed lines in the left panel represent average payments and true costs, respectively. See Appendix E for details.

As mentioned earlier, the decision-maker (min player) often does not have access to data such that $\pi(x, y') = 1$. Nevertheless, we will later show that it is possible to implement the above policy using data from $\pi(x, y') = 0$ (Section 4.3). We first prove that the proposed policy mitigates some of the effects observed under the status quo policy.

4.1 Revisiting the special case $y = y'$

Revisiting the special case introduced in the previous section, we can contrast the behavior of the status quo dynamics with that of our proposed policy reform. Whereas under the status quo both enrollment and costs increase without bound, under the alternative policy we observe the opposite behavior. Proofs of the statements below appear in Appendix C.4 and C.5.

Lemma 3. *Suppose the min player returns the canonical OLS solution in each round when the truncated design matrix no longer has full column rank. Then in every round t for which X_{π^t} is non-empty,*

$$\widehat{\mathbb{E}}\left[X_{\pi^t}\left(\theta^{t+1} - \theta^t\right)\right] := \frac{1}{n^t}\mathbf{1}^\top\left(X_{\pi^t}\left(\theta^{t+1} - \theta^t\right)\right) < 0\,,$$

where $n^t = |\{(x_i, y_i) : \pi^t(x_i, y_i) = 1, i \in [n]\}|$.

Proposition 2 (Equilibrium when $y = y'$ under (5))**.** *A fixed point of this game is:*

$$\pi^\star(x) \equiv 0 \qquad \textit{and} \qquad \theta^\star \textit{ is undefined}\,.$$

Phenomenon 6 (When not beneficial to the decision-maker, the program dies out). We remark that in this special case $y = y'$, the decision-maker does not ever benefit in members being served by the organization. Thus it is reasonable that the selection set becomes empty in this case, from the perspective of the decision-maker.

Phenomenon 7 (Using y' labels for the selected population is overly pessimistic). If we instead replace y with y' in Equation (5), the problem reduces to this special case (regardless of how y and y' relate). Thus, even with access to labels y' (which are often only privately known to the organizations), it may not be suitable to change the decision maker's objective in Equation (5) to use these realized costs y' as labels.

Using MA spending data to inform payment benchmarks has been proposed as a solution [12]. Our result suggests that estimating a payment model by fitting y' in the selected set would eliminate organizations' average margins at any nonempty equilibrium (by the OLS optimality condition), leaving organizations with no incentive to participate and potentially causing the program to die out.

4.2 General case

In the following, we consider the same problem setup as Section 3.2. We analyze the same quantities in Theorem 1 under the reformed policy in the following.

Theorem 2 (Reformed policy (see Theorem C.2))**.** *Suppose there exist $(\theta^\star, \theta_0^\star)$ and $\pi^\star$ which satisfy the fixed point equations. The following hold.*

$$\pi^\star(x, y) = \mathbb{1}\left\{x^\top\theta^\star + \theta_0^\star > y - \lambda(x)\right\} \textit{ and } (\theta^\star, \theta_0^\star) = \operatorname*{arg\,min}_{\theta, \theta_0} \mathbb{E}\left[(y - x^\top\theta - \theta_0)^2 \mid (x, y) \in \pi^\star\right]\,.$$

- ***No decision maker overpayment on the selected set.*** $\mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x, y') = 1\right] = 0.$

- ***Organization margin only comes from advantage due to*** $\lambda(x)$**.**

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y' \mid \pi^\star(x, y') = 1] = \mathbb{E}\left[\lambda(x) \mid \pi^\star(x, y') = 1\right].$$

- ***Selected population fraction.*** *Moreover, if* $|\alpha_0| < |\theta_0^\star - \beta_0|$, $\mathbb{E}[\pi^\star(x, y')] < 1/2$.

We illustrate these findings in Figure 4. The proof appears in Section C.1.2. Comparing Theorems 1 and 2, we make the following observations.

Phenomenon 8 (Decision maker has no overpayment under reformed policy). While the government overpayment under the status quo is positive, we see that it is 0 under the reformed policy (due to the optimality of $\theta^\star$ on $\pi^\star$ with respect to labels y).

Phenomenon 9 (Organization margin is only due to efficiency). We see that while the organization's margin formula is the same between both policies, the portion that is due to the real "efficiency" ($y - y' = \lambda(x)$) is the only contributing term to the margin under the reformed policy.

We cannot directly compare quantities between the two policies (which have different $\theta^\star$), but comparing margin expressions for the same $\theta^\star$ results in $\mathbb{E}[\lambda(x) \mid \pi^\star]$ being smaller under the status quo. This can be interpreted that for the same payment model, we "over-select" the population under the status quo. The organization profits for individuals with large $\lambda(x)$, i.e. $y' \ll y$, but selecting more individuals to benefit from government overpayment picks up lower valued $\lambda(x)$ as well.

Phenomenon 10 (Selected set is minority when average advantage is 0). Again considering $\alpha_0 = 0$, we see that the selected set has smaller mass under the reformed policy than under the status quo (less than half the population versus more than half). We may interpret this as more conservative selection of participants in the program, allowing the decision maker to eliminate overpayment while preserving organizations' margins from efficiency gains.

4.3 Implementing the Reformed Policy

Define estimators θ_{pop}, θ_π, $\theta_{\overline{\pi}}$ by $\theta_S := \arg\min_\theta \mathbb{E}\left[(y - x^\top\theta)^2 \mid (x, y) \in S\right]$ for S being the full population, π, or $\overline{\pi}$. Recall that in our setting, we only have access to $(x, y) \in \overline{\pi}$ and $(x, \emptyset) \in \pi$. Suppose we are able to solve for θ_{pop}, then we can recover θ_π. By first-order optimality of OLS, we have for each S: $\theta_S = \left(\mathbb{E}_x\left[xx^\top \mid (x, y) \in S\right]\right)^{-1} \mathbb{E}[xy \mid (x, y) \in S]$. The law of total expectation gives:

$$\begin{aligned}
\mathbb{E}_x\left[xx^\top\right] &= \Pr\left((x, y) \in \pi\right) \cdot \mathbb{E}_x\left[xx^\top \mid (x, y) \in \pi\right] + \Pr\left((x, y) \notin \pi\right) \cdot \mathbb{E}_x\left[xx^\top \mid (x, y) \notin \pi\right] \\
\mathbb{E}[xy] &= \Pr\left((x, y) \in \pi\right) \cdot \mathbb{E}[xy \mid (x, y) \in \pi] + \Pr\left((x, y) \notin \pi\right) \cdot \mathbb{E}[xy \mid (x, y) \notin \pi].
\end{aligned}$$

Suppose we can calculate $\mathbb{E}_x[xx^\top \mid (x, y) \in \pi]$. Let $p_\pi = \Pr\left((x, y) \in \pi\right)$. Then we can write θ_π in terms of the other two estimators as follows:

$$\theta_\pi = \left(\mathbb{E}_x[xx^\top \mid (x, y) \in \pi]\right)^{-1} \cdot \frac{1}{p_\pi}\left(\mathbb{E}_x[xx^\top] \cdot \theta_{pop} - (1 - p_\pi)\mathbb{E}_x[xx^\top \mid (x, y) \notin \pi] \cdot \theta_{\overline{\pi}}\right). \quad (6)$$

Estimating with access to $(x, y) \in \overline{\pi}$**.** When $x \sim \mathcal{N}(0, \Sigma)$, then θ_{pop} can be estimated with samples from $(x, y) \in \overline{\pi}$ without knowing $\lambda(x)$ using the method of [55, 56]. The procedure first estimates the set $\overline{\pi}$ from the available samples. Then, since (x, y) are jointly Gaussian, the projected SGD procedure of [55] recovers the true (population) Gaussian parameters, recovering θ_{pop}.

More generally, there are a number of heuristic methods which can be employed. For instance, one can instead impute the missing y labels in the selected set and fit OLS on the imputed labels.

5 Experiments

Experimental setup. We use the `upcoding` package to simulate a realistic distribution of co-occurring HCCs (diagnosis groups) in the Medicare population, using self-reported health conditions from a large national survey [61]. Each beneficiary is represented by a binary vector $x \in \{0, 1\}^d$, where $x_j = 1$ indicates the presence of HCC j, and we consider $n = 5000$ beneficiaries. We compute risk scores for the beneficiaries using the published CMS-HCC model coefficients, and annual costs are obtained by scaling these scores by the 2024 projected national per-capita FFS

Medicare spending benchmark of $13,261.20 [21]. Specifically, we set $y = (13,261.20) \cdot (\theta_0 + x^\top \theta^*) \cdot \exp(0.5\varepsilon - 0.125), \varepsilon \sim \mathcal{N}(0,1)$, where $\theta_0 = 0.35$ is the demographic baseline, and the multiplicative log-normal component captures the heavy right tail of healthcare spending [62]. To represent efficiency differences, we use $\lambda(x) = \alpha_0 + \alpha^\top x$, where $\alpha_0 > 0$ represents a baseline efficiency advantage and α assigns randomly sampled positive or negative cost advantages to a subset of clinical conditions to simulate heterogeneous service costs. Additional details on setup and data are deferred to Appendices E and F. Code to reproduce our experiments is available at this link.

Results. Figure 5 illustrates the dynamics under the status quo and reformed policies. Although the Gaussian assumptions used in our theoretical analysis do not hold in this setting, the qualitative behavior closely matches our results in Sections 3,4. Comparing the selection regions, we observe that under the status quo the insurer can profitably select both efficient and inefficient individuals, whereas under the reformed policy it selects only individuals for whom it is efficient. Further, the government overpayments are substantially larger under the status quo. In Figure 5 (bottom), enrollment rapidly increases under the status quo until nearly the entire population is selected, whereas under the reformed policy it decreases until overpayments are eliminated. The reform also selects individuals with a higher average efficiency advantage. Additional results are in Appendix F.

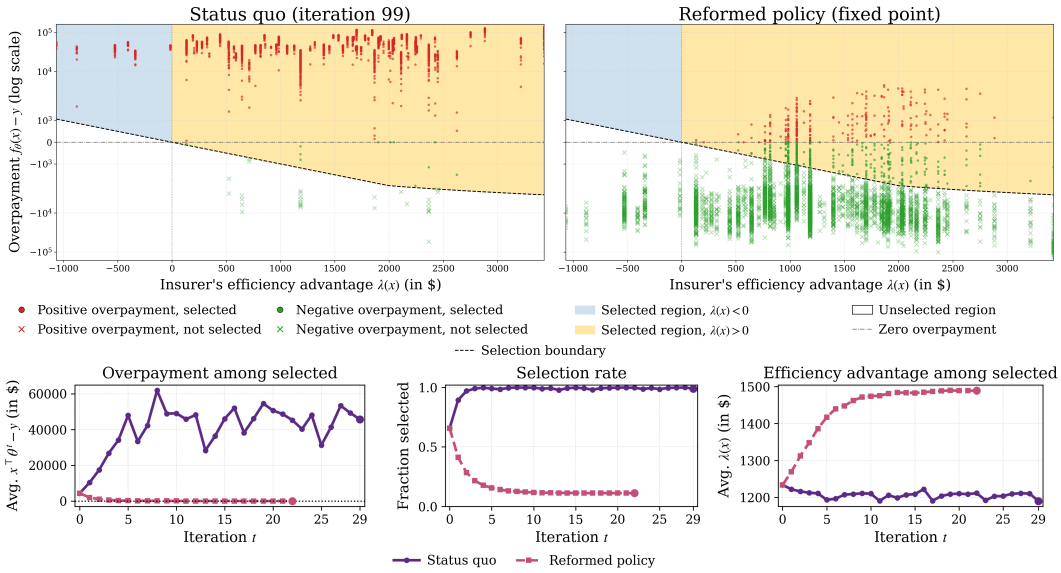

Figure 5: **Efficiency vs. Overpayment (Top).** Selection regions under the status quo at iteration 99 (top left) and the reformed policy at its fixed point (top right). **Dynamics (Bottom).** The bottom row shows the evolution of (1) average overpayment among selected beneficiaries, (2) selected population fraction $\mathbb{E}[\pi(x, y')]$, and (3) average efficiency advantage among selected beneficiaries.

6 Discussion and Future Work

We provide a theoretical framework characterizing the long-term behavior and design objectives of risk-based payment allocation. The empirical phenomena observed in practice emerge from studying the status quo dynamics in our framework, and we prove that our reformed policy mitigates these undesirable behaviors. While our experiments demonstrate this even when the idealized assumptions are violated, this is not fully explained by our theoretical analysis. Our proofs are based on analyzing the fixed points of the game, and it would be interesting to explore whether similar qualitative statements can be made for relaxed solution concepts. The implementation of our proposed policy uses known algorithms for regression with unknown truncation as a black-box, and therefore inherits their limitations. Thus, an interesting direction is to design algorithms that directly apply to the setting that we study in this paper with fewer additional assumptions. Beyond these considerations, reducing organizations' margins may have unintended consequences, such as changes in the quality of care or supplementary benefits offered to patients, which our framework does not capture.

Acknowledgments and Disclosure of Funding

We would like to thank Ahmed Alaa, Nika Haghtalab, Keegan Harris, Michael I. Jordan, Richard Kronick, Lydia T. Liu, Emma Pierson, Divya Shanmugam, Serena Wang, and Anna Zink for helpful conversations during the course of this project.

References

[1] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[2] Adriana Camacho and Emily Conover. Manipulation of social program eligibility. *American Economic Journal: Economic Policy*, 3(2):41–65, 2011.

[3] Rashiel B. Velarde. The Philippines' targeting system for the poor: Successes, lessons and ways forward. World Bank Social Protection Policy Note 16, World Bank, 2018.

[4] Sam Desiere, Kristine Langenbucher, and Ludo Struyven. Statistical profiling in public employment services: An international comparison. OECD Social, Employment and Migration Working Papers 224, OECD Publishing, 2019.

[5] Gregory C. Pope, John Kautter, Randall P. Ellis, Arlene S. Ash, John Z. Ayanian, Lisa I. Iezzoni, Melvin J. Ingber, Jesse M. Levy, and John Robst. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financing Review*, 25(4):119–141, 2004.

[6] Sherri Rose. A machine learning framework for plan payment risk adjustment. *Health Services Research*, 51(6):2358–2374, 2016.

[7] Marissa B. Reitsma, Thomas G. McGuire, and Sherri Rose. Algorithms to improve fairness in Medicare risk adjustment. *JAMA Health Forum*, 6(8):e252640, 2025.

[8] Medicare Payment Advisory Commission. The medicare advantage program: Status report. Report to the Congress: Medicare Payment Policy Chapter 12, Medicare Payment Advisory Commission (MedPAC), 2024.

[9] John Kautter, Gregory C. Pope, Melvin Ingber, Sara Freeman, Lindsey Patterson, Michael Cohen, and Patricia Keenan. The HHS-HCC risk adjustment model for individual and small group markets under the Affordable Care Act. *Medicare & Medicaid Research Review*, 4(3), 2014.

[10] Medicare Payment Advisory Commission. Report to the congress: Medicare payment policy. Technical report, Medicare Payment Advisory Commission, 2026.

[11] Medicare Payment Advisory Commission. Chapter 13: Estimating Medicare Advantage coding intensity and favorable selection. Technical report, Medicare Payment Advisory Commission, 2024.

[12] Medicare Payment Advisory Commission. Chapter 4: Favorable selection and future directions for Medicare Advantage payment policy. Technical report, Medicare Payment Advisory Commission, 2023.

[13] Medicare Payment Advisory Commission. March 2025 report to the Congress: Medicare payment policy. Technical report, Medicare Payment Advisory Commission (MedPAC), 2025.

[14] Jason Brown, Mark Duggan, Ilyana Kuziemko, and William Woolston. How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage program. *American Economic Review*, 104(10):3335–3364, 2014.

[15] Joseph P. Newhouse, Mary Price, Jie Huang, J. Michael McWilliams, and John Hsu. How much favorable selection is left in Medicare Advantage? *American Journal of Health Economics*, 1(1):1–26, 2015.

[16] Michael Geruso and Timothy Layton. Upcoding: Evidence from Medicare on squishy risk adjustment. *Journal of Political Economy*, 128(3):984–1026, 2020.

[17] Bowen Garrett. Favorable selection in Medicare Advantage? What the evidence tells us. Technical report, Urban Institute, 2024.

[18] Richard Kronick, F. Michael Chua, Ramona Krauss, Logan Johnson, and Daniel Waldo. Are fewer diagnoses better? Assessing a proposal to improve the Medicare Advantage payment system. *Health Affairs*, 44(1):66–74, 2025.

[19] Richard Kronick and F. Michael Chua. Industry-wide and sponsor-specific estimates of medicare advantage coding intensity. Technical report, Available at SSRN, November 2021.

[20] Nancy Ochieng, Meredith Freed, Jeannie Fuglesten Biniek, Anthony Damico, and Tricia Neuman. Medicare advantage in 2025: Enrollment update and key trends, July 2025.

[21] Centers for Medicare & Medicaid Services. Announcement of calendar year (CY) 2025 Medicare Advantage (MA) capitation rates and Part C and Part D payment policies. Technical report, Centers for Medicare & Medicaid Services, 2024.

[22] Sungchul Park, Anirban Basu, Norma Coe, and Fahad Khalil. Service-level selection: Strategic risk selection in Medicare Advantage in response to risk adjustment. Working Paper 24038, National Bureau of Economic Research, 2017.

[23] Richard Kronick. Better fit, same game? risk adjustment, coding incentives, and the limits of prediction. *Health Services Research*, 61(3):e70131, 2026.

[24] George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

[25] Michael Rothschild and Joseph Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

[26] Michael Spence. Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447, 1978.

[27] Mark Braverman and Sylvain Chassang. Data-driven incentive alignment in capitation schemes. Working Paper 28429, National Bureau of Economic Research, 2021.

[28] Donald B Gillies. Solutions to general non-zero-sum games. *Contributions to the Theory of Games*, 4(40):47–85, 1959.

[29] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.

[30] John C. Harsanyi. Games with incomplete information played by "bayesian" players, i-iii. part iii. the basic probability distribution of the game. *Management Science*, 14(7):486–502, 1968.

[31] Tao Li and Suresh P Sethi. A review of dynamic stackelberg game models. *Discrete & Continuous Dynamical Systems-Series B*, 22(1), 2017.

[32] Moritz Hardt, Nimrod Megiddo, Christos H. Papadimitriou, and Mary Wootters. Strategic classification. In *Proceedings of the ACM Conference on Innovations in Theoretical Computer Science*, 2016.

[33] John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[34] John Nash. Two-person cooperative games. *Econometrica*, 21(1):128–140, 1953.

[35] Robert J. Aumann. Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55(1):1–18, 1987.

[36] Sergiu Hart and David Schmeidler. Existence of correlated equilibria. *Mathematics of Operations Research*, 14(1):18–25, 1989.

[37] Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[38] Tim Roughgarden. Intrinsic robustness of the price of anarchy. *J. ACM*, 62(5), 2015.

[39] Marwaan Simaan and Jose B Cruz Jr. On the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(5):533–555, 1973.

[40] Philip J. Reny. On the existence of pure and mixed strategy nash equilibria in discontinuous games. *Econometrica*, 67(5):1029–1056, 1999.

[41] Francis Vella. Estimating models with sample selection bias: A survey. *The Journal of Human Resources*, 33(1):127–169, 1998.

[42] Patrick Puhani. The heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1):53–68, 2000.

[43] James Heckman. Varieties of selection bias. *The American Economic Review*, 80(2):313–318, 1990.

[44] G.S. Maddala. *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs. Cambridge University Press, 1983.

[45] Roderick J.A. Little and Donald B. Rubin. *Statistical Analysis with Missing Data*, volume 793. John Wiley & Sons, 2019.

[46] James J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

[47] Lung-Fei Lee. Some approaches to the correction of selectivity bias. *The Review of Economic Studies*, 49(3):355–372, 07 1982.

[48] Hyungtaik Ahn and James L. Powell. Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1):3–29, 1993.

[49] Mitali Das, Whitney K. Newey, and Francis Vella. Nonparametric estimation of sample selection models. *The Review of Economic Studies*, 70(1):33–58, 2003.

[50] A. Clifford Cohen. *Truncated and Censored Samples: Theory and Applications*. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 1991.

[51] Constantinos Daskalakis, Themis Gouleakis, Chistos Tzamos, and Manolis Zampetakis. Efficient Statistics, in High Dimensions, from Truncated Samples. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 639–649, 2018.

[52] Constantinos Daskalakis, Themis Gouleakis, Christos Tzamos, and Manolis Zampetakis. Computationally and Statistically Efficient Truncated Regression. In Alina Beygelzimer and Daniel Hsu, editors, *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99, pages 955–960, 2019.

[53] Constantinos Daskalakis, Dhruv Rohatgi, and Emmanouil Zampetakis. Truncated linear regression in high dimensions. In *Advances in Neural Information Processing Systems*, volume 33, pages 10338–10347, 2020.

[54] Andrew Ilyas, Emmanouil Zampetakis, and Constantinos Daskalakis. A theoretical and practical framework for regression and classification from truncated samples. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 2020.

[55] Jane H. Lee, Anay Mehrotra, and Manolis Zampetakis. Efficient statistics with unknown truncation, polynomial time algorithms, beyond gaussians. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 988–1006, 2024.

[56] Jane H Lee, Anay Mehrotra, and Manolis Zampetakis. Learning with positive and imperfect unlabeled data. *arXiv preprint arXiv:2504.10428*, 2025.

[57] Dennis Epple and Richard E. Romano. Competition between private and public schools, vouchers, and peer-group effects. *The American Economic Review*, 88(1):33–62, 1998.

[58] David Figlio and Joe Stone. Can public policy affect private school cream skimming? *Journal of Urban Economics*, 49(2):240–266, 2001.

[59] Rajashri Chakrabarti. Do vouchers lead to sorting under random private school selection? evidence from the milwaukee voucher program. *Economics of Education Review*, 34:191–218, 2013.

[60] Pierre Koning and Carolyn J. Heinrich. Cream-skimming, parking and other intended and unintended effects of high-powered, performance-based contracts. *Journal of Policy Analysis and Management*, 32(3):461–483, 2013.

[61] Oana M. Enache and Sherri Rose. Time-to-event estimation with unreliably reported events in medicare health plan payment. arXiv: 2602.04092, 2026.

[62] Eric French and John Bailey Jones. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics*, 19(6):705–721, 2004.

[63] Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the International Conference on Machine Learning*, 2020.

[64] Celestine Mendler-Dünner, Juan Perdomo, Tijana Zrnic, and Moritz Hardt. Stochastic optimization for performative prediction. In *Advances in Neural Information Processing Systems*, 2020.

[65] Himabindu Lakkaraju, Jon Kleinberg, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. The selective labels problem: Evaluating algorithmic predictions in the presence of unobservables. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

[66] Jennifer Chien, Margaret Roberts, and Berk Ustun. Algorithmic censoring in dynamic learning systems. In *Proceedings of the ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2023.

[67] Guy Horowitz, Yonatan Sommer, Moran Koren, and Nir Rosenfeld. Classification under strategic self-selection. In *Proceedings of the International Conference on Machine Learning*, 2024.

[68] Gregory Pope, John Kautter, Randall Ellis, Arlene Ash, John Ayanian, Lisa Lezzoni, Melvin Ingber, Jesse Levy, and John Robst. Risk adjustment of medicare capitation payments using the CMS-HCC model. *Health care financing review*, 25:119–41, 02 2004.

[69] Micah Johnson, Divya Shanmugam, David J. Meyers, Jenna Wiens, and Emma Pierson. The trillion-dollar algorithm: Lessons from machine learning for MA risk adjustment. *Health Affairs Forefront*, July 2026.

Appendix and Supplementary Material

Contents

1 Introduction **1**

1.1 Our Contributions . . . 2

1.2 Related Work . . . 3

2 Formal Framework **4**

3 Status Quo Policy **5**

3.1 Warm Up: Perfect Proxy y = y' . . . 5

3.2 General case . . . 6

4 Policy Reform **7**

4.1 Revisiting the special case y = y' . . . 8

4.2 General case . . . 8

4.3 Implementing the Reformed Policy . . . 9

5 Experiments **9**

6 Discussion and Future Work **10**

A Further Related Work **16**

B Mapping Design Choices to Real Allocation Systems **16**

C Proofs Deferred from Main Body **17**

C.1 Fixed Points Proofs . . . 18

C.1.1 Proof of Theorem 1 . . . 19

C.1.2 Proof of Theorem 2 . . . 22

C.1.3 Existence of Fixed Points . . . 25

C.2 Proof of Lemma 2 . . . 28

C.3 Proof of Proposition 1 . . . 29

C.4 Proof of Lemma 3 . . . 29

C.5 Proof of Proposition 2 . . . 30

D Additional Extensions of Main Results **30**

D.1 Extensions of Lemma 2 and Proposition 1 . . . 30

D.2 Beyond Gaussian Covariates . . . 31

D.3 Convergence Results . . . 33

D.4 Nonparametric Regression . . . 35

D.4.1 Sufficient Conditions . . . 35

D.4.2 Both y and y' depend on x linearly . . . 36

D.4.3 Explicit dependence as $y' = y - \lambda(x)$ 36

E Simulation Details **37**

F Experimental Details and Additional Experiments **37**

F.1 Data Description 37

F.2 Additional Experimental Details 37

F.3 Additional Results 39

A Further Related Work

In this section, we discuss how our work relates to prior work in machine learning on performative prediction, repeated risk minimization, selective labels, and strategic participation. We further elaborate on the relationship between our work and [27].

Performative prediction and repeated risk minimization. Let $S_\theta = \mathbf{1}\{Y' < f_\theta(X)\}$ denote the provider's best-response selection rule. After substituting this response, the status-quo update is repeated risk minimization on the model-dependent distribution $\mathcal{D}_\theta^- = \mathcal{L}(X, Y \mid S_\theta = 0)$, whereas the reform performs repeated risk minimization on $\mathcal{D}_\theta^+ = \mathcal{L}(X, Y \mid S_\theta = 1)$; their fixed points with positive selected and unselected population mass can therefore be viewed as performatively stable points [63, 64]. Standard existence and convergence tools from this literature may apply at the population level under suitable regularity and distributional sensitivity assumptions. Our local contraction analysis (Appendix D.3) gives sufficient conditions for the update map in our setting. The principal distinction with our work is twofold: (1) we show how our distribution map is generated by a profit-maximizing provider in the Medicare setting, and most importantly, (2) the existence of stable points is an initial observation for us; most of our work focuses on analyzing the properties of stable points and designing mechanisms that have stable points with some specified properties. In particular, our focus is on the economic properties of these fixed points, including government overpayment, organizational margins, and enrollment. We characterize these properties in a risk-based allocation setting and study how changing the fitting objective affects them.

Selective labels and strategic participation. Our data-access problem is related to selective labels, where outcomes are observed only following one side of a decision [65], as well as to dynamic censoring and strategic participation [66, 67]. In our setting, however, the label-revealing action is chosen by the provider using private costs, the selection boundary is determined by the current payment rule and can therefore change after each retraining round, and the model explicitly distinguishes the planner's counterfactual cost y from the provider's realized cost y'. These works primarily target predictive performance and evaluation, fairness, or the effects of selective observation and participation, in some cases by introducing exploration or otherwise changing which labels are observed. Our reform instead targets equilibrium transfers and reconstructs the selected-population counterfactual regression using the truncation structure, without requiring randomized service assignment.

Risk selection and capitation design. Closest in application, [27] explain why improving ex ante risk prediction alone need not eliminate favorable selection and propose an ex post approach to incentive alignment. Their focus is capitation-mechanism design rather than feedback generated by repeatedly refitting the payment model on an endogenously retained population. Our analysis is therefore complementary: we isolate this learning-system mechanism, characterize its fixed-point implications for both overpayment and enrollment, and propose a selected-population counterfactual fitting rule that can be implemented under the planner's existing label-access constraints (Section 4.3).

B Mapping Design Choices to Real Allocation Systems

Decision maker's objectives. The decision maker in our setting is the government/public payer rather than a profit-maximizing entity. It cares about the overall cost of providing services: delegation is beneficial when the organization can provide services more efficiently than direct public provision,

provided that the government does not pay more than its own counterfactual cost of provision. In particular, our framework asks whether delegation generates genuine efficiency gains without requiring the public payer to spend more than it would have under direct provision.

These goals are captured by the two quantities defined in our framework. For the selected population,

$$\mathbb{E}[y - y' \mid \pi = 1] = \mathbb{E}[f_\theta(x) - y' \mid \pi = 1] - \mathbb{E}[f_\theta(x) - y \mid \pi = 1].$$

Thus, even if the government remains cost-neutral, delegation is beneficial when $y' < y$ because services are provided at a lower overall cost, and the resulting margin gives the organization an incentive to participate. On the other hand, when $y = y'$, there are no genuine efficiency gains from services being provisioned by the organization, and organizational margin comes entirely from government overpayment.

Rebates. Our model abstracts from how financial gains are shared among the organization, the public payer, and beneficiaries. In Medicare Advantage, payments are determined through a benchmark, bid, and rebate system. The benchmark is based on the projected local spending under traditional FFS Medicare and incorporates statutory adjustments and quality bonuses. When an insurance plan bids below the benchmark, a share of the difference is paid to the plan as a rebate, while CMS retains the remainder. Rebate dollars are used to provide supplemental benefits or lower beneficiary costs through lower premiums or cost sharing, with allowances for associated administrative expenses and profit. Thus, in practice, the savings do not accrue entirely to the organization.

We can incorporate such sharing through a stylized extension. Assume that the organization retains a fixed fraction $\alpha \in (0, 1]$ of its payment–cost margin, yielding a payoff of

$$\alpha\big(f_\theta(x) - y'\big).$$

Since $\alpha > 0$, its selection rule remains $f_\theta(x) > y'$, and hence the retraining and selection dynamics are unchanged. This extension shows that fixed proportional sharing leaves the dynamics unchanged, although it does not capture the full Medicare Advantage rebate system. We use a simpler formulation to isolate the selection and retraining dynamics.

Cost/Information asymmetry. The decision maker observes costs y for individuals who remain under its own policy ($\pi = 0$). We assume that organizations also have access to these observations, whereas their own costs y' are private and not observed by the decision maker. This reflects the Medicare Advantage setting we study, where reported encounter and financial data do not necessarily reveal the complete beneficiary-level realized cost of serving each enrollee.

The organizations can thus use information about y' to select individuals while the decision maker fits the payment model using only the costs y it observes from the $\pi = 0$ population. For individuals selected by the organization ($\pi = 1$), the counterfactual cost y that would have been incurred under the status quo policy is observed by neither the decision maker nor the organization.

In practice, organizations may have richer information about costs than is captured by the variables used in the public payer's risk-adjustment model. The public payer sets payments using the restricted variables in the risk-adjustment model. On the other hand, the organization can use additional utilization, clinical, and proprietary information not included in the payment rule to estimate its costs.

Another source of asymmetry between publicly usable and privately available information is the organization's ability to use more complex model classes, whereas interpretability is an important consideration in policy settings. For example, CMS uses a linear model and restricts it to clinically meaningful diagnosis categories [68]. Recent work in health policy also notes that Medicare Advantage plans are able to preferentially enroll less costly patients based on factors not captured in the risk model, such as social determinants of health or individual preferences for care, whereas risk scores explain only approximately 14% of the individual variation in medical spending [69].

Here, y' can also be interpreted as the organization's privately estimated conditional expected cost in the selection rule, and our results will continue to hold under this interpretation.

C Proofs Deferred from Main Body

In this section, we include full statements and proofs of the results presented in the main text.

C.1 Fixed Points Proofs

Lemma C.1 (Fixed point quantities)**.** *Suppose $x \sim \mathcal{N}(0, \Sigma)$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon \perp x$ and we have the relation $y' = y - \lambda(x)$, where $\lambda(x) = x^\top \alpha + \alpha_0$. Let $\theta^\star$ be arbitrary and*

$$\pi^\star(x, y) = \mathbb{1}\left\{x^\top \theta^\star + \theta_0^\star > y - \lambda(x)\right\}.$$

Define

$$\kappa(a) \coloneqq \frac{\phi(a)}{\Phi(a)}, \;\; a^\star = \frac{\theta_0^\star - \beta_0 + \alpha_0}{\sqrt{\sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma (\theta^\star - \beta + \alpha)}}, \;\; s^\star = \sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma (\theta^\star - \beta + \alpha).$$

Then the following hold:

$$\mathbb{E}\left[(\theta_0^\star + x^\top \theta^\star) - y \mid \pi^\star(x, y') = 1\right] = (\theta_0^\star - \beta_0) + \kappa(a^\star) \cdot \frac{(\theta^\star - \beta)^\top \Sigma (\theta^\star - \beta + \alpha) + \sigma^2}{\sqrt{s^\star}}$$
$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y' \mid \pi^\star(x, y') = 1] = (a^\star + \kappa(a^\star)) \cdot \sqrt{s^\star} > 0$$
$$\mathbb{E}[\pi^\star(x, y')] = \Phi(a^\star)$$

Proof. We assume the design matrix has a column of $\mathbf{1}$ (i.e., the insurer runs linear regression with an intercept). Write

$$y = \beta_0 + x^\top \beta + \varepsilon, \qquad \lambda(x) = \alpha_0 + \alpha^\top x.$$

Selection occurs when when $\theta_0^\star + x^\top \theta^\star > y - \lambda(x)$, which is equivalent to

$$\theta_0^\star + x^\top \theta^\star > \beta_0 + x^\top \beta + \varepsilon - (\alpha_0 + x^\top \alpha) \iff \varepsilon < (\theta_0^\star - \beta_0 + \alpha_0) + x^\top (\theta^\star - \beta + \alpha).$$

This implies for $\pi^\star(x, y') = 1 \iff \varepsilon - x^\top (\theta^\star - \beta + \alpha) < \theta_0^\star - \beta_0 + \alpha_0$..

Define $z \coloneqq \varepsilon - x^\top (\theta^\star - \beta + \alpha)$. Since (x, ε) is jointly Gaussian, z is Gaussian with mean 0 and variance

$$s^\star = \sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma (\theta^\star - \beta + \alpha).$$

Hence

$$z \sim \mathcal{N}(0, s^\star), \qquad a^\star = \frac{\theta_0^\star - \beta_0 + \alpha_0}{\sqrt{s^\star}},$$

and

$$\pi^\star(x, y') = 1 \iff z < \theta_0^\star - \beta_0 + \alpha_0$$

We will use the standard Gaussian truncation identities: if $z \sim \mathcal{N}(0, s^\star)$ and $t = \theta_0^\star - \beta_0 + \alpha_0$, then

$$\Pr(z < t) = \Phi\left(\frac{t}{\sqrt{s^\star}}\right) = \Phi(a^\star), \qquad \mathbb{E}[z \mid z < t] = -\kappa(a^\star)\sqrt{s^\star}.$$

We will also use the standard conditional Gaussian formula for the jointly Gaussian pair (x, z):

$$\mathbb{E}[x \mid z = u] = \frac{\mathrm{Cov}(x, z)}{\mathrm{Var}(z)}\, u.$$

Now

$$\mathrm{Cov}(x, z) = \mathrm{Cov}\big(x, \varepsilon - x^\top (\theta^\star - \beta + \alpha)\big) = -\Sigma(\theta^\star - \beta + \alpha),$$

since $\varepsilon \perp x$. Therefore

$$\mathbb{E}[x \mid z = u] = -\frac{\Sigma(\theta^\star - \beta + \alpha)}{s^\star}\, u.$$

Taking expectation over the event $z < t$ and using $\mathbb{E}[z \mid z < t] = -\kappa(a^\star)\sqrt{s^\star}$ gives

$$\mathbb{E}[x \mid \pi^\star(x, y') = 1] = \frac{\kappa(a^\star)}{\sqrt{s^\star}} \Sigma(\theta^\star - \beta + \alpha).$$

Decision maker's overpayment on the selected population. Using $y = \beta_0 + x^\top \beta + \varepsilon$,

$$(\theta_0^\star + x^\top \theta^\star) - y = (\theta_0^\star - \beta_0) + x^\top(\theta^\star - \beta) - \varepsilon.$$

Also, from the definition of z,

$$\varepsilon = z + x^\top(\theta^\star - \beta + \alpha).$$

Hence

$$(\theta_0^\star + x^\top \theta^\star) - y = (\theta_0^\star - \beta_0) - z - x^\top \alpha.$$

Taking conditional expectation on $\{\pi^\star(x, y') = 1\} = \{z < t\}$,

$$\begin{aligned}\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y \mid \pi^\star(x, y') = 1] &= (\theta_0^\star - \beta_0) - \mathbb{E}[z \mid z < t] - \mathbb{E}[x \mid z < t]^\top \alpha \\ &= (\theta_0^\star - \beta_0) + \kappa(a^\star)\sqrt{s^\star} - \frac{\kappa(a^\star)}{\sqrt{s^\star}}(\theta^\star - \beta + \alpha)^\top \Sigma \alpha.\end{aligned}$$

Since

$$s^\star = \sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha),$$

the last two terms simplify to

$$\kappa(a^\star) \cdot \frac{\sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta)}{\sqrt{s^\star}}.$$

Therefore

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y \mid \pi^\star(x, y') = 1] = (\theta_0^\star - \beta_0) + \kappa(a^\star) \cdot \frac{(\theta^\star - \beta)^\top \Sigma(\theta^\star - \beta + \alpha) + \sigma^2}{\sqrt{s^\star}}.$$

Insurer's margin on the selected population. Since $y' = y - \lambda(x)$,

$$\begin{aligned}(\theta_0^\star + x^\top \theta^\star) - y' &= (\theta_0^\star + x^\top \theta^\star) - y + \lambda(x) \\ &= (\theta_0^\star - \beta_0 + \alpha_0) + x^\top(\theta^\star - \beta + \alpha) - \varepsilon \\ &= (\theta_0^\star - \beta_0 + \alpha_0) - z.\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y' \mid \pi^\star(x, y') = 1] &= (\theta_0^\star - \beta_0 + \alpha_0) - \mathbb{E}[z \mid z < t] \\ &= (\theta_0^\star - \beta_0 + \alpha_0) + \kappa(a^\star)\sqrt{s^\star}.\end{aligned}$$

Since

$$\theta_0^\star - \beta_0 + \alpha_0 = a^\star \sqrt{s^\star},$$

we obtain

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y' \mid \pi^\star(x, y') = 1] = (a^\star + \kappa(a^\star))\sqrt{s^\star}.$$

Fraction of the selected population. Finally,

$$\mathbb{E}[\pi^\star(x, y')] = \Pr(\pi^\star(x, y') = 1) = \Pr(z < \theta_0^\star - \beta_0 + \alpha_0) = \Phi(a^\star).$$

□

C.1.1 Proof of Theorem 1

Theorem C.1 (Status quo)**.** *Suppose there exist $\theta^\star$ and $\pi^\star$ which satisfy the fixed point equations.* $\pi^\star(x, y) = \mathbb{1}\left\{x^\top \theta^\star + \theta_0^\star > y - \lambda(x)\right\}$ *and* $(\theta^\star, \theta_0^\star) = \arg\min_{\theta, \theta_0} \mathbb{E}\left[(y - x^\top \theta - \theta_0)^2 \mid (x, y) \in \overline{\pi^\star}\right]$.

Define

$$\rho(a) \coloneqq \frac{\phi(a)}{1 - \Phi(a)}, \quad \tau(a) \coloneqq \rho(a)(\rho(a) - a),$$

$$c^\star = \frac{\tau(a^\star)\sigma^2}{\sigma^2 + \big(1 - \tau(a^\star)\big)(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)},$$

and

$$M^\star = \sigma^2 + c^\star(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha).$$

Then the quantity $\theta_0^\star - \beta_0$ satisfies

$$\theta_0^\star - \beta_0 = \frac{\rho(a^\star)}{\sqrt{s^\star}} \cdot M^\star > 0.$$

In particular, we can write the following.

Decision-maker overpayment on selected set is positive.

$$\mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1\right] = \frac{\rho(a^\star)+\kappa(a^\star)}{\sqrt{s^\star}} \cdot M^\star > 0\,.$$

Organization margin on selected comes from both overpayment and advantage due to $\lambda(x)$.

$$\begin{aligned}\mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y' \mid \pi^\star(x,y') = 1] &= \mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1\right] + \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right] \\ &= \frac{\rho(a^\star)+\kappa(a^\star)}{\sqrt{s^\star}} \cdot M^\star + \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right] \\ \Rightarrow \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right] &= (a^\star + \kappa(a^\star)) \cdot \sqrt{s^\star} - \frac{\rho(a^\star)+\kappa(a^\star)}{\sqrt{s^\star}} \cdot M^\star\end{aligned}$$

Selected population size. *Moreover,*

$$a^\star = \frac{\theta_0^\star - \beta_0 + \alpha_0}{\sqrt{s^\star}} = \frac{\theta_0^\star - \beta_0}{\sqrt{s^\star}} + \frac{\alpha_0}{\sqrt{s^\star}}\,.$$

Hence, if $|\alpha_0| < \theta_0^\star - \beta_0$, then $a^\star > 0$, and therefore, $\mathbb{E}[\pi^\star(x,y')] = \Phi(a^\star) > 1/2$.

Proof. Because $(\theta_0^\star, \theta^\star)$ is a fixed point of the status quo policy, it satisfies the population first-order conditions for OLS on the non-selected population:

$$\mathbb{E}\left[\begin{pmatrix}1\\ x\end{pmatrix}\left(y - \theta_0^\star - x^\top\theta^\star\right) \,\middle|\, \pi^\star(x,y') = 0\right] = 0.$$

As in the proof of Lemma C.1, define $z := \varepsilon - x^\top(\theta^\star - \beta + \alpha)$. Then

$$\pi^\star(x,y') = 0 \iff z \geq \theta_0^\star - \beta_0 + \alpha_0.$$

We use the Gaussian truncation identities for the upper tail. In particular,

$$\mathbb{E}[x \mid \pi^\star(x,y') = 0] = -\frac{\rho(a^\star)}{\sqrt{s^\star}}\Sigma(\theta^\star - \beta + \alpha), \quad \mathbb{E}[\varepsilon \mid \pi^\star(x,y') = 0] = \frac{\rho(a^\star)}{\sqrt{s^\star}}\sigma^2,$$

$$\mathbb{E}[xx^\top \mid \pi^\star(x,y') = 0] = \Sigma + \frac{a^\star\rho(a^\star)}{s^\star}\Sigma(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top\Sigma,$$

and

$$\mathbb{E}[x\varepsilon \mid \pi^\star(x,y') = 0] = -\frac{a^\star\rho(a^\star)\sigma^2}{s^\star}\Sigma(\theta^\star - \beta + \alpha).$$

Since $y - \theta_0^\star - x^\top\theta^\star = \varepsilon - (\theta_0^\star - \beta_0) - x^\top(\theta^\star - \beta)$, the intercept first-order condition gives

$$\mathbb{E}[\varepsilon \mid \pi^\star = 0] - (\theta_0^\star - \beta_0) - \mathbb{E}[x \mid \pi^\star = 0]^\top(\theta^\star - \beta) = 0.$$

Substituting the truncated-moment formulas above,

$$\frac{\rho(a^\star)}{\sqrt{s^\star}}\sigma^2 - (\theta_0^\star - \beta_0) + \frac{\rho(a^\star)}{\sqrt{s^\star}}(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta) = 0.$$

Hence

$$\theta_0^\star - \beta_0 = \frac{\rho(a^\star)}{\sqrt{s^\star}}\Big(\sigma^2 + (\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta)\Big). \tag{7}$$

Next, the slope first-order condition is

$$\mathbb{E}[x\varepsilon \mid \pi^\star = 0] - (\theta_0^\star - \beta_0)\mathbb{E}[x \mid \pi^\star = 0] - \mathbb{E}[xx^\top \mid \pi^\star = 0](\theta^\star - \beta) = 0.$$

Substituting the same identities,

$$\begin{aligned}&-\frac{a^\star\rho(a^\star)\sigma^2}{s^\star}\Sigma(\theta^\star - \beta + \alpha) + \frac{\rho(a^\star)}{\sqrt{s^\star}}(\theta_0^\star - \beta_0)\Sigma(\theta^\star - \beta + \alpha) \\ &\qquad - \left(\Sigma + \frac{a^\star\rho(a^\star)}{s^\star}\Sigma(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top\Sigma\right)(\theta^\star - \beta) = 0.\end{aligned}$$

Rearranging,

$$\left(I + \frac{a^\star \rho(a^\star)}{s^\star}(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top \Sigma\right)(\theta^\star - \beta) = \left(-\frac{a^\star \rho(a^\star)\sigma^2}{s^\star} + \frac{\rho(a^\star)}{\sqrt{s^\star}}(\theta_0^\star - \beta_0)\right)(\theta^\star - \beta + \alpha). \quad (8)$$

Thus $\theta^\star - \beta$ is collinear with $\theta^\star - \beta + \alpha$, so there exists a scalar $c^\star$ such that

$$\theta^\star - \beta = c^\star(\theta^\star - \beta + \alpha).$$

Computing $c^\star$. Substituting $\theta^\star - \beta = c^\star(\theta^\star - \beta + \alpha)$ into (7) gives

$$\theta_0^\star - \beta_0 = \frac{\rho(a^\star)}{\sqrt{s^\star}}\Big(\sigma^2 + c^\star(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)\Big). \quad (9)$$

Now, substituting both $\theta^\star - \beta$ and $\theta_0^\star - \beta_0$ into (8), the left-hand side becomes

$$c^\star\left(1 + \frac{a^\star \rho(a^\star)}{s^\star}(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)\right)(\theta^\star - \beta + \alpha),$$

and the right-hand side becomes,

$$\left(-\frac{a^\star \rho(a^\star)\sigma^2}{s^\star} + \frac{\rho(a^\star)^2}{s^\star}\Big(\sigma^2 + c^\star(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)\Big)\right)(\theta^\star - \beta + \alpha).$$

Equating the scalar coefficients of $(\theta^\star - \beta + \alpha)$ and rearranging yields

$$c^\star\Big(s^\star + a^\star \rho(a^\star)(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha) - \rho(a^\star)^2(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)\Big) = \rho(a^\star)(\rho(a^\star) - a^\star)\sigma^2. \quad (10)$$

By definition of $\tau(a^\star) = \rho(a^\star)(\rho(a^\star) - a^\star)$ and $s^\star = \sigma^2 + (\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)$,

$$c^\star = \frac{\tau(a^\star)\sigma^2}{\sigma^2 + \big(1 - \tau(a^\star)\big)(\theta^\star - \beta + \alpha)^\top \Sigma(\theta^\star - \beta + \alpha)}.$$

Using the definition of $M^\star$, equation (7) becomes

$$\theta_0^\star - \beta_0 = \frac{\rho(a^\star)}{\sqrt{s^\star}} M^\star.$$

Since $\rho(a^\star) > 0$ and $M^\star > 0$, this implies

$$\theta_0^\star - \beta_0 > 0.$$

By Lemma C.1,

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y \mid \pi^\star(x, y') = 1] = (\theta_0^\star - \beta_0) + \kappa(a^\star)\frac{M^\star}{\sqrt{s^\star}}.$$

Substituting the expression for $\theta_0^\star - \beta_0$,

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y \mid \pi^\star(x, y') = 1] = \frac{\rho(a^\star) + \kappa(a^\star)}{\sqrt{s^\star}} M^\star > 0.$$

Also by Lemma C.1,

$$\mathbb{E}[(\theta_0^\star + x^\top \theta^\star) - y' \mid \pi^\star(x, y') = 1] = (a^\star + \kappa(a^\star))\sqrt{s^\star}.$$

Since $y' = y - \lambda(x)$,

$$(\theta_0^\star + x^\top \theta^\star) - y' = \big((\theta_0^\star + x^\top \theta^\star) - y\big) + \lambda(x),$$

and therefore

$$\mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y' \mid \pi^\star(x,y') = 1] = \mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1] + \mathbb{E}[\lambda(x) \mid \pi^\star(x,y') = 1].$$

Combining the previous two equations,

$$\mathbb{E}[\lambda(x) \mid \pi^\star(x,y') = 1] = (a^\star + \kappa(a^\star))\sqrt{s^\star} - \frac{\rho(a^\star) + \kappa(a^\star)}{\sqrt{s^\star}} M^\star.$$

Finally,

$$a^\star = \frac{\theta_0^\star - \beta_0 + \alpha_0}{\sqrt{s^\star}} = \frac{\theta_0^\star - \beta_0}{\sqrt{s^\star}} + \frac{\alpha_0}{\sqrt{s^\star}}.$$

If $|\alpha_0| < \theta_0^\star - \beta_0$, then

$$\theta_0^\star - \beta_0 + \alpha_0 > 0,$$

hence $a^\star > 0$. Since Φ is strictly increasing and $\Phi(0) = 1/2$,

$$\mathbb{E}[\pi^\star(x,y')] = \Phi(a^\star) > \frac{1}{2}.$$

□

C.1.2 Proof of Theorem 2

Theorem C.2 (Reformed policy)**.** *Suppose there exist $\theta^\star$ and $\pi^\star$ which satisfy the fixed point equations.*

$$\pi^\star(x,y) = \mathbb{1}\left\{x^\top\theta^\star + \theta_0^\star > y - \lambda(x)\right\} \text{ and } (\theta^\star, \theta_0^\star) = \arg\min_{\theta,\theta_0} \mathbb{E}\left[(y - x^\top\theta - \theta_0)^2 \mid (x,y) \in \pi^\star\right].$$

Define

$$\widetilde{\tau}(a) \coloneqq \kappa(a)\big(\kappa(a) + a\big), \qquad \widetilde{c}^\star = \frac{\widetilde{\tau}(a^\star)\sigma^2}{\sigma^2 + \big(1 - \widetilde{\tau}(a^\star)\big)(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta + \alpha)},$$

and

$$\widetilde{M}^\star = \sigma^2 + \widetilde{c}^\star(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta + \alpha).$$

Then the quantity $\theta_0^\star - \beta_0$ satisfies

$$\theta_0^\star - \beta_0 = -\frac{\kappa(a^\star)}{\sqrt{s^\star}} \cdot \widetilde{M}^\star < 0.$$

In particular, we can write the following.

Decision maker has no overpayment on the selected set.

$$\mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1\right] = 0\,.$$

Organization margin only comes from advantage due to $\lambda(x)$.

$$\mathbb{E}[(\theta_0^\star + x^\top\theta^\star) - y' \mid \pi^\star(x,y') = 1] = \mathbb{E}\left[(\theta_0^\star + x^\top\theta^\star) - y \mid \pi^\star(x,y') = 1\right] + \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right]$$
$$\Rightarrow \mathbb{E}\left[\lambda(x) \mid \pi^\star(x,y') = 1\right] = (a^\star + \kappa(a^\star)) \cdot \sqrt{s^\star}\,.$$

Selected population size. *Moreover,*

$$a^\star = \frac{\theta_0^\star - \beta_0 + \alpha_0}{\sqrt{s^\star}} = \frac{\theta_0^\star - \beta_0}{\sqrt{s^\star}} + \frac{\alpha_0}{\sqrt{s^\star}}\,.$$

Hence, if $|\alpha_0| < |\theta_0^\star - \beta_0|$, then $a^\star < 0$, and therefore, $\mathbb{E}[\pi^\star(x,y')] = \Phi(a^\star) < 1/2$.

Proof. Because $(\theta_0^\star, \theta^\star)$ is a fixed point of the reformed policy, it satisfies the population first-order conditions for OLS on the selected population:

$$\mathbb{E}\left[\begin{pmatrix}1\\ x\end{pmatrix}(y - \theta_0^\star - x^\top\theta^\star) \;\middle|\; \pi^\star(x, y') = 1\right] = 0.$$

As in the proof of Lemma C.1, define $z := \varepsilon - x^\top(\theta^\star - \beta + \alpha)$. Then

$$\pi^\star(x, y') = 1 \iff z < \theta_0^\star - \beta_0 + \alpha_0.$$

We use the Gaussian truncation identities for the lower tail. In particular,

$$\mathbb{E}[x \mid \pi^\star(x, y') = 1] = \frac{\kappa(a^\star)}{\sqrt{s^\star}}\Sigma(\theta^\star - \beta + \alpha), \quad \mathbb{E}[\varepsilon \mid \pi^\star(x, y') = 1] = -\frac{\kappa(a^\star)}{\sqrt{s^\star}}\sigma^2,$$

$$\mathbb{E}[xx^\top \mid \pi^\star(x, y') = 1] = \Sigma - \frac{a^\star\kappa(a^\star)}{s^\star}\Sigma(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top\Sigma,$$

and

$$\mathbb{E}[x\varepsilon \mid \pi^\star(x, y') = 1] = \frac{a^\star\kappa(a^\star)\sigma^2}{s^\star}\Sigma(\theta^\star - \beta + \alpha).$$

Since $y - \theta_0^\star - x^\top\theta^\star = \varepsilon - (\theta_0^\star - \beta_0) - x^\top(\theta^\star - \beta)$, the intercept first-order condition gives

$$\mathbb{E}[\varepsilon \mid \pi^\star = 1] - (\theta_0^\star - \beta_0) - \mathbb{E}[x \mid \pi^\star = 1]^\top(\theta^\star - \beta) = 0.$$

Substituting the truncated-moment formulas above,

$$-\frac{\kappa(a^\star)}{\sqrt{s^\star}}\sigma^2 - (\theta_0^\star - \beta_0) - \frac{\kappa(a^\star)}{\sqrt{s^\star}}(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta) = 0.$$

Hence

$$\theta_0^\star - \beta_0 = -\frac{\kappa(a^\star)}{\sqrt{s^\star}}\Big(\sigma^2 + (\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta)\Big). \tag{11}$$

Next, the slope first-order condition is

$$\mathbb{E}[x\varepsilon \mid \pi^\star = 1] - (\theta_0^\star - \beta_0)\mathbb{E}[x \mid \pi^\star = 1] - \mathbb{E}[xx^\top \mid \pi^\star = 1](\theta^\star - \beta) = 0.$$

Substituting the same identities,

$$\begin{aligned}\frac{a^\star\kappa(a^\star)\sigma^2}{s^\star}\Sigma(\theta^\star - \beta + \alpha) - \frac{\kappa(a^\star)}{\sqrt{s^\star}}(\theta_0^\star - \beta_0)\Sigma(\theta^\star - \beta + \alpha)\\ - \left(\Sigma - \frac{a^\star\kappa(a^\star)}{s^\star}\Sigma(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top\Sigma\right)(\theta^\star - \beta) = 0.\end{aligned}$$

Rearranging,

$$\begin{aligned}\left(I - \frac{a^\star\kappa(a^\star)}{s^\star}(\theta^\star - \beta + \alpha)(\theta^\star - \beta + \alpha)^\top\Sigma\right)(\theta^\star - \beta) =\\ \left(\frac{a^\star\kappa(a^\star)\sigma^2}{s^\star} - \frac{\kappa(a^\star)}{\sqrt{s^\star}}(\theta_0^\star - \beta_0)\right)(\theta^\star - \beta + \alpha).\end{aligned} \tag{12}$$

Thus $\theta^\star - \beta$ is collinear with $\theta^\star - \beta + \alpha$, so there exists a scalar $\widetilde{c}^\star$ such that

$$\theta^\star - \beta = \widetilde{c}^\star(\theta^\star - \beta + \alpha).$$

Computing $\widetilde{c}^\star$. Substituting $\theta^\star - \beta = \widetilde{c}^\star(\theta^\star - \beta + \alpha)$ into (11) gives

$$\theta_0^\star - \beta_0 = -\frac{\kappa(a^\star)}{\sqrt{s^\star}}\Big(\sigma^2 + \widetilde{c}^\star(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta + \alpha)\Big). \tag{13}$$

Now, substituting both $\theta^\star - \beta$ and $\theta_0^\star - \beta_0$ into (12), the left-hand side becomes

$$\widetilde{c}^\star\left(1 - \frac{a^\star\kappa(a^\star)}{s^\star}(\theta^\star - \beta + \alpha)^\top\Sigma(\theta^\star - \beta + \alpha)\right)(\theta^\star - \beta + \alpha),$$

and the right-hand side becomes

$$\left(\frac{a^\star\kappa(a^\star)\sigma^2}{s^\star}+\frac{\kappa(a^\star)^2}{s^\star}\Big(\sigma^2+\widetilde{c}^\star(\theta^\star-\beta+\alpha)^\top\Sigma(\theta^\star-\beta+\alpha)\Big)\right)(\theta^\star-\beta+\alpha).$$

Equating the scalar coefficients of $(\theta^\star-\beta+\alpha)$ and rearranging yields

$$\widetilde{c}^\star\Big(s^\star-a^\star\kappa(a^\star)(\theta^\star-\beta+\alpha)^\top\Sigma(\theta^\star-\beta+\alpha)-\kappa(a^\star)^2(\theta^\star-\beta+\alpha)^\top\Sigma(\theta^\star-\beta+\alpha)\Big)= \kappa(a^\star)(\kappa(a^\star)+a^\star)\sigma^2. \quad (14)$$

By definition of $\widetilde{\tau}(a^\star)=\kappa(a^\star)(\kappa(a^\star)+a^\star)$ and $s^\star=\sigma^2+(\theta^\star-\beta+\alpha)^\top\Sigma(\theta^\star-\beta+\alpha)$,

$$\widetilde{c}^\star=\frac{\widetilde{\tau}(a^\star)\sigma^2}{\sigma^2+\big(1-\widetilde{\tau}(a^\star)\big)(\theta^\star-\beta+\alpha)^\top\Sigma(\theta^\star-\beta+\alpha)}.$$

Using the definition of $\widetilde{M}^\star$, equation (11) becomes

$$\theta_0^\star-\beta_0=-\frac{\kappa(a^\star)}{\sqrt{s^\star}}\widetilde{M}^\star.$$

Since $\kappa(a^\star)>0$ and $\widetilde{M}^\star>0$, this implies

$$\theta_0^\star-\beta_0<0.$$

By the OLS first-order condition on the selected population,

$$\mathbb{E}[(\theta_0^\star+x^\top\theta^\star)-y\mid\pi^\star(x,y')=1]=0.$$

Also by Lemma C.1,

$$\mathbb{E}[(\theta_0^\star+x^\top\theta^\star)-y'\mid\pi^\star(x,y')=1]=(a^\star+\kappa(a^\star))\sqrt{s^\star}.$$

Since $y'=y-\lambda(x)$,

$$(\theta_0^\star+x^\top\theta^\star)-y'=\big((\theta_0^\star+x^\top\theta^\star)-y\big)+\lambda(x),$$

and therefore

$$\begin{aligned}\mathbb{E}[(\theta_0^\star+x^\top\theta^\star)-y'\mid\pi^\star(x,y')=1]&=\mathbb{E}[(\theta_0^\star+x^\top\theta^\star)-y\mid\pi^\star(x,y')=1]+\mathbb{E}[\lambda(x)\mid\pi^\star(x,y')=1]\\&=0+\mathbb{E}[\lambda(x)\mid\pi^\star(x,y')=1].\end{aligned}$$

Combining the previous two equations,

$$\mathbb{E}[\lambda(x)\mid\pi^\star(x,y')=1]=(a^\star+\kappa(a^\star))\sqrt{s^\star}.$$

Finally,

$$a^\star=\frac{\theta_0^\star-\beta_0+\alpha_0}{\sqrt{s^\star}}=\frac{\theta_0^\star-\beta_0}{\sqrt{s^\star}}+\frac{\alpha_0}{\sqrt{s^\star}}.$$

If $|\alpha_0|<|\theta_0^\star-\beta_0|$, then since $\theta_0^\star-\beta_0<0$,

$$\theta_0^\star-\beta_0+\alpha_0<0,$$

hence $a^\star<0$. Since Φ is strictly increasing and $\Phi(0)=1/2$,

$$\mathbb{E}[\pi^\star(x,y')]=\Phi(a^\star)<\frac{1}{2}.$$

□

C.1.3 Existence of Fixed Points

Lemma C.2 (Existence)**.** *Consider the setting of Theorem C.1. If $\alpha \not\equiv 0$, a fixed point exists for every $\alpha_0 \in \mathbb{R}$. If $\alpha \equiv 0$, a finite fixed point exists if and only if $\alpha_0 < 0$.*

Proof. Recall the following:

$$\kappa(a) \coloneqq \frac{\phi(a)}{\Phi(a)}\,, \quad \rho(a) \coloneqq \frac{\phi(a)}{1-\Phi(a)}\,, \quad \tau(a) \coloneqq \rho(a)\left(\rho(a)-a\right)\,.$$

For any candidate coefficients (θ, θ_0) define

$$s \coloneqq \sigma^2 + (\theta-\beta+\alpha)^\top \Sigma\,(\theta-\beta+\alpha)\,, \quad a \coloneqq \frac{\theta_0-\beta_0+\alpha_0}{\sqrt{s}}\,,$$

$$c \coloneqq \frac{\tau(a)\sigma^2}{\sigma^2 + (1-\tau(a))\,(\theta-\beta+\alpha)^\top \Sigma\,(\theta-\beta+\alpha)}\,, \quad M \coloneqq \sigma^2 + c\,(\theta-\beta+\alpha)^\top \Sigma\,(\theta-\beta+\alpha)\,.$$

where a superscript $(\star)$ will denote these quantities at a fixed point.

Step 1: Fixed point conditions on slope and intercept. Recall as in the proof of Theorem 1, the Gaussian truncation identities yield the following first-order conditions for the slope and intercept, respectively:

$$\theta-\beta = c\,(\theta-\beta+\alpha) \qquad \text{Slope Condition}$$

$$\theta_0-\beta_0 = \frac{\rho(a)}{\sqrt{s}}\left(\sigma^2 + (\theta-\beta+\alpha)^\top \Sigma\,(\theta-\beta)\right) \qquad \text{Intercept Condition}$$

Since the retained covariate distribution is absolutely continuous, distinct affine predictors differ with positive conditional probability. Hence the conditional squared-loss objective is strictly convex in (θ, θ_0), and its minimizer is unique.

Substituting the slope condition into the intercept condition gives

$$\theta_0-\beta_0 = \frac{\rho(a)}{\sqrt{s}} M\,.$$

Step 2: Solve slope condition for each possible threshold. First, assume that $\alpha \not\equiv 0$. Starting from the slope condition,

$$\begin{aligned}
\theta-\beta &= c\,(\theta-\beta+\alpha) \\
&= c\,(\theta-\beta) + c\alpha \\
\Rightarrow (1-c)\,(\theta-\beta) &= c\alpha \\
\Rightarrow \theta-\beta &= \frac{c}{1-c}\alpha \qquad \text{since } 0<c<1 \\
\Rightarrow \theta-\beta+\alpha &= \left(\frac{c}{1-c}\alpha + \alpha\right) \\
&= \frac{\alpha}{1-c} \\
\Rightarrow (\theta-\beta+\alpha)^\top \Sigma\,(\theta-\beta+\alpha) &= \frac{1}{(1-c)^2}\cdot \alpha^\top \Sigma \alpha\,.
\end{aligned}$$

Substituting into the definition of c,

$$\begin{aligned}
c &= \frac{\tau(a)\cdot\sigma^2}{\sigma^2 + \frac{1-\tau(a)}{(1-c)^2}\alpha^\top\Sigma\alpha} \\
\Rightarrow c\left(\sigma^2 + \frac{1-\tau(a)}{(1-c)^2}\alpha^\top\Sigma\alpha\right) &= \tau(a)\sigma^2 \\
\Rightarrow c\frac{1-\tau(a)}{(1-c)^2}\alpha^\top\Sigma\alpha &= (\tau(a)-c)\,\sigma^2 \\
\Rightarrow \frac{c}{\tau(a)-c} &= \frac{\sigma^2(1-c)^2}{(1-\tau(a))\alpha^\top\Sigma\alpha}\,.
\end{aligned}$$

where in the last line we've used that $1-\tau(a)>0, \alpha^\top\Sigma\alpha>0, c>0 \Rightarrow \tau(a)-c>0$. For each fixed a, we will solve the above equation on $0<c<\tau(a)$. The left-hand side has derivative

$$\frac{d}{dc}\left(\frac{c}{\tau(a)-c}\right)=\frac{\tau(a)}{(\tau(a)-c)^2}>0\,.$$

The right-hand side has derivative

$$\frac{d}{dc}\frac{\sigma^2(1-c)^2}{(1-\tau(a))\alpha^\top\Sigma\alpha}=-\frac{2\sigma^2(1-c)}{(1-\tau(a))\alpha^\top\Sigma\alpha}<0\,.$$

As c approaches 0, the left-hand side tends to 0 while the right-hand side approaches a positive number. As c approaches $\tau(a)$, the left-hand side diverges to infinity while the right-hand side remains finite. By monotonicity, the two sides must cross exactly once at some $c\in(0,\tau(a))$. For each fixed a, then there is a unique solution $c(a)\in(0,\tau(a))$.

Step 3: Solve intercept condition. For each $a\in\mathbb{R}$, the previous step determines $c(a)$ and

$$\theta(a)=\beta+\frac{c(a)}{1-c(a)}\alpha\,.$$

Evaluating the definitions of s and M at $c(a)$, define:

$$s(a)\coloneqq\sigma^2+\frac{\alpha^\top\Sigma\alpha}{(1-c(a))^2}\quad\text{and}\quad M(a)\coloneqq\sigma^2+c(a)\frac{\alpha^\top\Sigma\alpha}{(1-c(a))^2}\,.$$

The intercept condition requires

$$\theta_0-\beta_0=\frac{\rho(a)}{\sqrt{s(a)}}M(a)\,.$$

Substituting this into the definition of a:

$$a=\rho(a)\frac{M(a)}{s(a)}+\frac{\alpha_0}{\sqrt{s(a)}}\,. \tag{15}$$

The following also holds from the definition of $c(a)$:

$$\begin{aligned}
c(a)&=\frac{\tau(a)\sigma^2}{\sigma^2+\frac{1-\tau(a)}{(1-c(a))^2}\alpha^\top\Sigma\alpha}\\
\Rightarrow c(a)\left(\sigma^2+\frac{1-\tau(a)}{(1-c(a))^2}\alpha^\top\Sigma\alpha\right)&=\tau(a)\sigma^2\\
\Rightarrow c(a)\sigma^2+c(a)\frac{1}{(1-c(a))^2}\alpha^\top\Sigma\alpha&=\tau(a)\sigma^2+\tau(a)c(a)\frac{1}{(1-c(a))^2}\alpha^\top\Sigma\alpha\\
\Rightarrow c(a)\left(\sigma^2+\frac{1}{(1-c(a))^2}\alpha^\top\Sigma\alpha\right)&=\tau(a)\left(\sigma^2+\frac{c(a)}{(1-c(a))^2}\alpha^\top\Sigma\alpha\right)\\
\Rightarrow c(a)s(a)&=\tau(a)M(a)\\
\Rightarrow\frac{M(a)}{s(a)}&=\frac{c(a)}{\tau(a)}\,.
\end{aligned}$$

Thus Equation (15) is equivalent to

$$a=\frac{\rho(a)c(a)}{\tau(a)}+\frac{\alpha_0}{\sqrt{s(a)}}\,.$$

Define the function

$$F(a)\coloneqq a-\frac{\rho(a)c(a)}{\tau(a)}-\frac{\alpha_0}{\sqrt{s(a)}}\,.$$

This function is continuous, since $c(a), s(a), \rho(a), \tau(a)$ are continuous, with $s(a)>0, \tau(a)>0$. It remains to show that $F(a)=0$ for some a to find a fixed point.

Step 4: Show that $F(a)$ has a zero. Since $F(a)$ is a continuous function, and $\rho(a) - a > 0$, we can multiply the expression for $F(a)$ with strictly positive factors and will show that

$$H(a) \coloneqq \sqrt{s(a)}\,(\rho(a) - a)\,F(a) = \sqrt{s(a)}\,(a(\rho(a) - a) - c(a)) - \alpha_0(\rho(a) - a)\,,$$

changes sign in order to show that it (and thus $F(a)$) has a zero.

Case $a \to \infty$. As $a \to \infty$, it holds that

$$\rho(a) - a \to 0\,, \quad 1 - \tau(a) \to 0\,, \quad \text{and} \quad \frac{(\rho(a) - a)^2}{1 - \tau(a)} \to 1\,,$$

which is done by repeated integration by parts of Gaussian tail quantities and using a Taylor expansion to get an expansion of these functions in terms of orders of a to evaluate these limits. The following also holds from the definition of $c(a)$:

$$(1 - c(a))^3 = (1 - \tau(a)) \left(\alpha^\top \Sigma \alpha \cdot \frac{c(a)}{\sigma^2} + (1 - c(a))^2 \right)\,,$$

where since $0 < c(a) < 1$,

$$0 \leq \alpha^\top \Sigma \alpha \cdot \frac{c(a)}{\sigma^2} + (1 - c(a))^2 \leq \frac{1}{\sigma^2} \alpha^\top \Sigma + 1 \quad \Rightarrow \quad (1 - c(a))^3 \to 0\,,$$

thus $c(a) \to 1$ using the fact that $1 - \tau(a) \to 0$ to evaluate the right-hand side limit in the identity. Dividing the identity yields

$$\frac{1 - \tau(a)}{1 - c(a)} = \frac{(1 - c(a))^2}{\frac{c(a)}{\sigma^2} \alpha^\top \Sigma \alpha + (1 - c(a))^2} \to 0 \quad \Rightarrow \quad \frac{(\rho(a) - a)^2}{1 - c(a)} \to 0\,,$$

by using the fact that $\frac{(\rho(a) - a)^2}{1 - \tau(a)} \to 1$.

Finally,

$$H(a) = \sqrt{\sigma^2 (1 - c(a))^2 + \alpha^\top \Sigma \alpha} \left(1 - \frac{1 - \tau(a)}{1 - c(a)} - \frac{(\rho(a) - a)^2}{1 - c(a)} \right) - \alpha_0 (\rho(a) - a)\,.$$

The square root term tends to $\alpha^\top \Sigma \alpha$, the middle term tends to 1, and the last term tends to 0. Thus $H(a) \to \sqrt{\alpha^\top \Sigma \alpha}$ as $a \to \infty$.

Case $a \to -\infty$. As $a \to -\infty$, it holds that

$$c(a) \to 0\,, \quad s(a) \to \sigma^2 + \alpha^\top \Sigma \alpha\,, \quad \text{and} \quad \rho(a) \to 0\,.$$

Then,

$$\frac{H(a)}{a^2} \to -\sqrt{\sigma^2 + \alpha^\top \Sigma \alpha} < 0\,,$$

thus $H(a) \to -\infty$ as $a \to -\infty$.

Thus by using the intermediate value theorem, $H(a)$ (and thus $F(a)$) has a finite zero at some $a^\star$ with $H(a^\star) = 0 = F(a^\star)$.

Step 5: Construct fixed point. We've established that there exists $a^\star \in \mathbb{R}$ such that $F(a^\star) = 0$. Set

$$c^\star = c(a^\star)\,, \quad s^\star = s(a^\star)\,, \quad \text{and} \quad M^\star = M(a^\star)\,,$$

and define

$$\theta^\star = \beta + \frac{c^\star}{1 - c^\star} \alpha\,, \quad \text{and} \quad \theta_0^\star = \beta_0 + \frac{\rho(a^\star)}{\sqrt{s^\star}} M^\star\,.$$

Step 2 ensures the slope fixed-point equation and step 3 ensures the intercept fixed-point condition. Step 1 shows that these coefficients are exactly the unique conditional least-squares minimizer.

Step 6: Handle $\alpha \equiv 0$. Suppose $\alpha \equiv 0$. Then $c < 1$ and the slope condition forces

$$\theta - \beta = c(\theta - \beta) \quad \Rightarrow \quad \theta = \beta\,.$$

Then this also results in $s = \sigma^2$, $M = \sigma^2$, $c = \tau(a)$, and the intercept condition yields

$$\theta_0 - \beta_0 = \sigma\rho(a)\,.$$

By the definition of a:

$$a = \frac{\theta_0 - \beta_0 + \alpha_0}{\sqrt{s}} = \rho(a) + \frac{\alpha_0}{\sigma}\,,$$

it must hold that $a - \rho(a) = \frac{\alpha_0}{\sigma}$. Differentiating with respect to a, $\rho'(a) = \tau(a)$ and thus $\frac{d}{da}(a - \rho(a) = 1 - \tau(a)$. So $a - \rho(a)$ is continuous and strictly increasing. Also, using $\rho(a) \to 0$ as $a \to -\infty$ and $\rho(a) - a \to 0$ as $a \to \infty$ from earlier, it holds that $a - \rho(a) \to -\infty$ as $a \to -\infty$ and $a - \rho(a) \to 0$ as $a \to \infty$. Since $\rho(a) > a$ for all a, the range of $a - \rho(a)$ is $(-\infty, 0)$. Thus the fixed point equations have a solution only if $\alpha_0 < 0$.

□

Corollary C.1 (Reform Policy)**.** *Consider the setting of Theorem 2. If $\alpha \not\equiv 0$, a fixed point exists for every $\alpha_0 \in \mathbb{R}$. If $\alpha \equiv 0$, a finite fixed point exists if and only if $\alpha_0 > 0$.*

Proof. Apply the preceding result under the change of variables

$$(y, \beta, \beta_0, \varepsilon, \alpha, \alpha_0, \theta, \theta_0) \mapsto (-y, -\beta, -\beta_0, -\varepsilon, -\alpha, -\alpha_0, -\theta, -\theta_0)\,.$$

The Gaussian assumptions and squared loss are unchanged, while the selected event

$$y < x^\top\theta + \theta_0 + \lambda(x)$$

becomes the retained event, up to a probability-zero boundary. Under this transformation, $a \mapsto -a$ and $s \mapsto s$, with

$$\rho(-a) = \kappa(a), \qquad \tau(-a) = \widetilde{\tau}(a)\,.$$

Consequently, the original quantities c and M become $\widetilde{c}$ and $\widetilde{M}$, and the intercept identity becomes

$$\theta_0^\star - \beta_0 = -\frac{\kappa(a^\star)}{\sqrt{s^\star}}\widetilde{M}^\star.$$

Thus the result follows, giving existence for every α_0 when $\alpha \neq 0$, and existence exactly when $-\alpha_0 < 0$, or $\alpha_0 > 0$, when $\alpha = 0$. □

C.2 Proof of Lemma 2

Proof. Fix some round t. The max player chooses

$$\pi^t(x) = \mathbb{1}\left\{(\theta^t)^\top x > y\right\}\,.$$

Then the min player responds with

$$\theta^{t+1} = X^+_{\overline{\pi^t}} y_{\overline{\pi^t}}\,.$$

By construction, we have that

$$X_{\overline{\pi^t}}\theta^t \leq y_{\overline{\pi^t}}\,.$$

Then

$$\begin{aligned} X_{\overline{\pi^t}}\left(\theta^{t+1} - \theta^t\right) &= X_{\overline{\pi^t}} X^+_{\overline{\pi^t}} y_{\overline{\pi^t}} - X_{\overline{\pi^t}}\theta^t \\ &\geq X_{\overline{\pi^t}} X^+_{\overline{\pi^t}} y_{\overline{\pi^t}} - y_{\overline{\pi^t}} && \text{(using } X_{\overline{\pi^t}}\theta^t \leq y_{\overline{\pi^t}}\text{)} \\ &= \left(X_{\overline{\pi^t}} X^+_{\overline{\pi^t}} - I\right) y_{\overline{\pi^t}} \end{aligned}$$

Note that when $X_{\overline{\pi^t}}$ has full row rank, this is equal to 0 immediately.

Let $P := X_{\overline{\pi^t}} X^+_{\overline{\pi^t}}$, the projection matrix which maps vectors to the column space of $X_{\overline{\pi^t}}$. Then

$$X_{\overline{\pi^t}}(\theta^{t+1} - \theta^t) \geq (P - I) y_{\overline{\pi^t}}\,.$$

Taking empirical averages,

$$\widehat{\mathbb{E}}\big[X_{\overline{\pi^t}}(\theta^{t+1}-\theta^t)\big] \geq \frac{1}{\left|\overline{\pi^t}\right|}\mathbf{1}^\top(P-I)y_{\overline{\pi^t}}.$$

If the design matrix contains a constant column, then $\mathbf{1} \in \mathrm{col}(X_{\overline{\pi^t}})$, so $P\mathbf{1} = \mathbf{1}$ and hence

$$\mathbf{1}^\top(P-I)y_{\overline{\pi^t}} = 0.$$

Therefore,

$$\widehat{\mathbb{E}}\big[X_{\overline{\pi^t}}(\theta^{t+1}-\theta^t)\big] \geq 0.$$

The only inequality was due to $X_{\overline{\pi^t}}\theta^t \leq y_{\overline{\pi^t}}$, thus we can replace it by equality when this condition holds with equality. □

C.3 Proof of Proposition 1

Proof. Let $\overline{\pi^\star} = \{x \in \mathcal{X} : \pi^\star(x) = 0\}$ denote the set of points not selected by the max player. By definition of $\pi^\star$, this set satisfies

$$\overline{\pi^\star} = \{x \in \mathcal{X} : x^\top\theta^\star \leq y\}.$$

We now consider two cases.

Case 1 (Fixed point is reached). Suppose there exists a non-empty set $S \subseteq \mathcal{X}$ and a vector $\widetilde{\theta} = X_s^+ y_S$ such that

$$x^\top\widetilde{\theta} = y \text{ for all } x \in S, \qquad x^\top\widetilde{\theta} > y \text{ for all } x \notin S.$$

Define $\pi^\star(x) = 1 - \mathbb{1}_S(x)$, so that $\overline{\pi^\star} = S$. By construction,

$$x^\top\widetilde{\theta} = y \text{ for all } x \in S$$

so $\widetilde{\theta}$ is the (minimum norm) solution to the min player's objective (so $\theta^\star = \widetilde{\theta}$). In one more step, since all points outside S satisfy $x^\top\widetilde{\theta} > y$, the max player's best response is

$$\pi^\star(x) = \mathbb{1}\{x^\top\widetilde{\theta} > y\} = 1 - \mathbb{1}_S(x).$$

The max player remains at $\pi^\star(x) = 1 - \mathbb{1}_S(x)$ and the min player again chooses $\theta^\star = \widetilde{\theta}$.

Case 2. Suppose no such non-empty set S exists. Suppose toward a contradiction that the game reaches equilibrium at round t. In the next round, we have

$$\pi^t(x) = \mathbb{1}\left\{x^\top\theta^t > y\right\}.$$

There must exist some x such that

$$x^\top\theta^t < y\,,$$

with strict inequality, otherwise we would end up in Case 1. Then the set defined by $\overline{\pi^{t+1}}$ is non-empty. By the earlier lemma,

$$\widehat{\mathbb{E}}_{x\,:\,\pi^{t+1}(x)=0}\left[x^\top\left(\theta^{t+1}-\theta^t\right)\right] > 0\,,$$

holds with strict inequality since $x^\top\theta^t < y$. Thus $\theta^{t+1} \neq \theta^t$. This is a contradiction. □

C.4 Proof of Lemma 3

Proof. Fix some round t. Again, the max player chooses

$$\pi^t(x) = \mathbb{1}\left\{x^\top\theta^t > y\right\}.$$

The min player responds with

$$\theta^{t+1} = X_{\pi^t}^+ y_{\pi^t}\,.$$

By construction, we have

$$X_{\pi^t}\theta^t > y_{\pi^t}\,.$$

Then

$$\begin{aligned} X_{\pi^t}\left(\theta^{t+1}-\theta^t\right) &= X_{\pi^t}X_{\pi^t}^+ y_{\pi^t} - X_{\pi^t}\theta^t \\ &< X_{\pi^t}X_{\pi^t}^+ y_{\pi^t} - y_{\pi^t} && \text{(using } X_{\pi^t}\theta^t > y_{\pi^t}\text{)} \\ &= \left(X_{\pi^t}X_{\pi^t}^+ - I\right) y_{\pi^t}\,. \end{aligned}$$

Note that above is $(P-I)y_{\pi^t}$ for $P = X_{\pi^t}X_{\pi^t}^+$. By the same arguments in Lemma 2, we can conclude that

$$\widehat{\mathbb{E}}\left[X_{\pi^t}\left(\theta^{t+1}-\theta^t\right)\right] < 0\,.$$

□

C.5 Proof of Proposition 2

Proof. Suppose toward a contradiction that at equilibrium in round t, there exists some x such that $\pi^t(x) = 1$. Then X_{π^t} is non-empty, and by Lemma 3, we have:

$$\widehat{\mathbb{E}}\left[X_{\pi^t}\left(\theta^{t+1}-\theta^t\right)\right] < 0\,.$$

In particular $\theta^{t+1} \neq \theta^t$, which is a contradiction.

When $\pi^\star(x) \equiv 0$, then no points are selected. When no points are selected, the min player's solution $\theta^\star$ is undefined. In the next round, since $\theta^\star$ is undefined, the max player does not select any points and remains at $\pi^\star$. The min player remains at $\theta^\star$. □

D Additional Extensions of Main Results

D.1 Extensions of Lemma 2 and Proposition 1

In this section we write the population versions of the results we presented for the special case $y = y'$ in the main body. We also consider $\lambda(x)$ here with $\lambda(x) \equiv 0$ as the special case $y = y'$.

Lemma D.1 (One-step increase)**.** *Consider the dynamics described by Equation (4) for $t = 1, 2, \ldots$. It holds that*

$$\mathbb{E}\left[x^\top\left(\theta^{t+1}-\theta^t\right) \mid \overline{\pi^t}\right] \geq \mathbb{E}\left[\lambda(x) \mid \overline{\pi^t}\right]\,,$$

where equality holds iff $y - x^\top\theta = \lambda(x)$ a.s on $\overline{\pi^t}$.

Proof. Fix a round t and parameter θ^t. Then π^t is described by

$$\pi^t(x,y) = \mathbb{1}\left\{x^\top\theta^t > y - \lambda(x)\right\} \iff \overline{\pi^t}(x,y) = \mathbb{1}\left\{x^\top\theta^t \leq y - \lambda(x)\right\} = \mathbb{1}\left\{y - x^\top\theta^t \geq \lambda(x)\right\}\,.$$

The next θ^{t+1}, by the first-order condition, satisfies

$$\mathbb{E}\left[y - x^\top\theta^{t+1} \mid \overline{\pi^t}\right] = 0 \Rightarrow \mathbb{E}\left[x^\top\theta^{t+1} \mid \overline{\pi^t}\right] = \mathbb{E}\left[y \mid \overline{\pi^t}\right]\,.$$

Then we can write:

$$\begin{aligned} \mathbb{E}\left[x^\top\left(\theta^{t+1}-\theta^t\right)\right] &= \mathbb{E}\left[y - x^\top\theta^t \mid \overline{\pi^t}\right] && \text{(using } \mathbb{E}\left[x^\top\theta^{t+1} \mid \overline{\pi^t}\right] = \mathbb{E}\left[y \mid \overline{\pi^t}\right]\text{)} \\ &\geq \mathbb{E}\left[\lambda(x) \mid \overline{\pi^t}\right]\,. && \text{(using } y - x^\top\theta^t \geq \lambda(x) \text{ in } \overline{\pi^t}\text{)} \end{aligned}$$

Note that the only inequality used was due to $y - x^\top\theta^t \geq \lambda(x)$ in $\overline{\pi^t}$, thus equality holds iff $y - x^\top\theta^t = \lambda(x)$ almost surely in $\overline{\pi^t}$. □

Corollary D.1. *When $\lambda(x) \equiv 0$ (i.e., $y = y'$) or more generally when $\lambda(x) \geq 0$ for all x, then $\mathbb{E}\left[x^\top\left(\theta^{t+1}-\theta^t\right) \mid \overline{\pi^t}\right] \geq 0$.*

Remark 5 ($\lambda(x) \geq 0$ makes more aggressive progress)**.** We can extend Lemma 2 for $\lambda(x) \geq 0$ for all x to again conclude Proposition 1. In this case, we see that the one-step increase is not only lower bounded by 0 but potentially a positive number, which more aggressively moves the model upward. In practice, we can also see this behavior, e.g., if we always set $\lambda(x) = \lambda > 0$. Even when $\lambda < 0$, we see the same behavior but slower progress. See Figure 6.

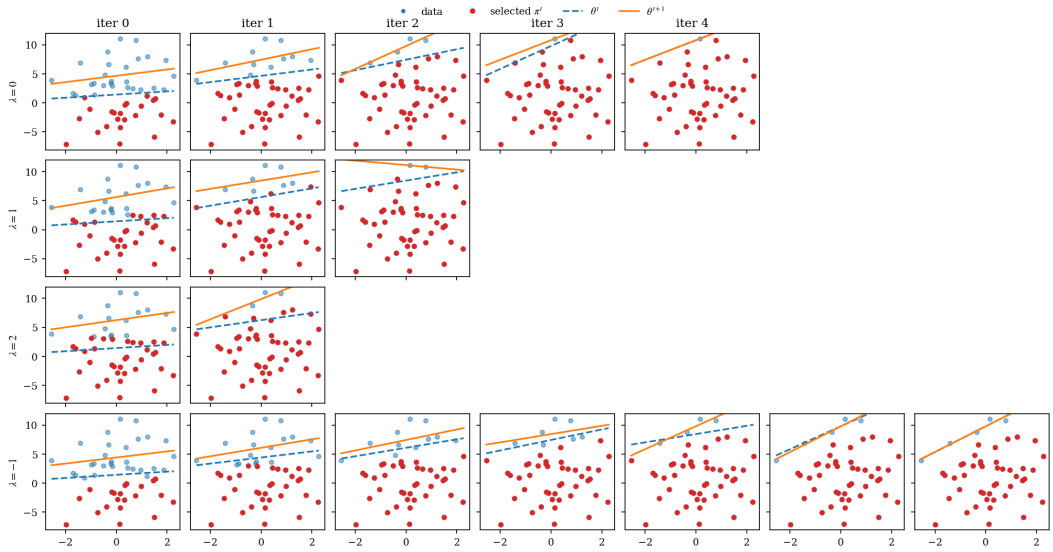

Figure 6: Dynamics of Equation (4) but $\lambda(x) \equiv \lambda$ for some constant λ.

Proposition D.1 (Fixed points are boundary equilibria)**.** *Consider the dynamics described by Equation (4) when $y = y'$. Let $\theta^\star$ and $\pi^\star$ satisfy the fixed points equations under Equation (4). Suppose* $\Pr\left(\overline{\pi^\star}\right) > 0$ *and*

$$\theta^\star \in \arg\min_\theta \mathbb{E}\left[(y - x^\top\theta)^2 \mid (x,y) \in \overline{\pi^\star}\right] .$$

Then

$$y = x^\top\theta^\star \text{ a.s. on } \overline{\pi^\star} .$$

Proof. By the optimality of $\theta^\star$ on $\overline{\pi^\star}$:

$$\mathbb{E}[y - x^\top\theta^\star \mid \overline{\pi^\star}] = 0 .$$

But by definition

$$y - x^\top\theta^\star \geq 0 \qquad \text{for all } (x,y) \in \overline{\pi^\star} .$$

Since a non-negative random variable with expectation 0 is 0 almost surely, we get that $y = x^\top\theta^\star$ almost surely on $\overline{\pi^\star}$. □

Remark 6. This population statement says that the mass on $\overline{\pi^\star}$ must all collapse onto a line. In finite samples, this was a set with some number of points which must all lie on a line. In the population version, we may not expect such a line to have positive mass in general.

Corollary D.2 (Nonexistence of nontrivial population fixed points)**.** *Suppose there does not exist $\theta^\star, \pi^\star$ satisfying the fixed point equations in Equation (4) such that*

$$\Pr(\overline{\pi^\star}) > 0 \qquad \textit{and} \qquad y = x^\top\theta^\star \text{ a.s. on } \overline{\pi^\star} .$$

Then there are no fixed points with positive retained mass. In particular,

$$\pi^\star \equiv 1 \qquad \textit{and} \qquad \theta^\star \textit{ is undefined} ,$$

is such fixed point with no positive retained mass.

D.2 Beyond Gaussian Covariates

One of our key main results shows that at equilibrium, the government overpayment is positive under the status quo policy, regardless of the parameter $\theta^\star$. The other properties (e.g., statements about the organization's margins) follow from this or optimality conditions. While assuming Gaussianity

allows us to write concrete expressions which are clearly positive, we can reach a similar conclusion under weaker conditions. The core property we need is

$$\mathbb{E}\left[x^\top\theta^\star - y \mid x^\top\theta^\star - y + \lambda(x) > 0\right] > 0\,.$$

for fixed points $\pi^\star$ and $\theta^\star$. Recall the relationship between the overpayment and the selection rule:

$$r_\theta(x,y) = x^\top\theta - y, \qquad x^\top\theta - y + \lambda(x) = r_\theta(x,y) + \lambda(x) \quad \Rightarrow \quad \pi_\theta(x,y) = \mathbb{1}\left\{r_\theta(x,y) + \lambda(x) > 0\right\}\,.$$

Then we can write

$$\begin{aligned}
\mathrm{Cov}\left(r_{\theta^\star}(x,y), \pi_{\theta^\star}(x,y)\right) &= \mathbb{E}\left[r_{\theta^\star}\mathbb{1}\left\{r_{\theta^\star} + \lambda(x) > 0\right\}\right] - \mathbb{E}\left[r_{\theta^\star}\right]\cdot\mathbb{E}\left[\mathbb{1}\left\{r_{\theta^\star} + \lambda(x) > 0\right\}\right] \\
\Rightarrow \mathbb{E}\left[r_{\theta^\star}\mathbb{1}\left\{r_{\theta^\star} + \lambda(x) > 0\right\}\right] &= \mathrm{Cov}\left(r_{\theta^\star}, \pi_{\theta^\star}\right) + \mathbb{E}\left[r_{\theta^\star}\right]\cdot\mathbb{E}\left[\mathbb{1}\left\{r_{\theta^\star} + \lambda(x) > 0\right\}\right] \\
\Rightarrow \mathbb{E}\left[r_{\theta^\star}(x,y) \mid \pi_{\theta^\star}(x,y)\right] &= \mathbb{E}[r_{\theta^\star}(x,y)] + \frac{\mathrm{Cov}\left(r_{\theta^\star}(x,y), \mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right)}{\mathbb{E}\left[\mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right]}\,.
\end{aligned}$$

The condition we need is

$$\mathbb{E}[r_{\theta^\star}(x,y)] + \frac{\mathrm{Cov}\left(r_{\theta^\star}(x,y), \mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right)}{\mathbb{E}\left[\mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right]} > 0$$

$$\Rightarrow \mathrm{Cov}\left(r_{\theta^\star}(x,y), \mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right) > -\mathbb{E}[r_{\theta^\star}(x,y)]\cdot\Pr\left(r_{\theta^\star}(x,y) + \lambda(x) > 0\right) \quad (16)$$

If $\mathbb{E}[r_{\theta^\star}(x,y)] \geq 0$, this is asking for the overpayments and the selection rule to be correlated/associated (i.e., $\mathrm{Cov}\left(r_{\theta^\star}(x,y), \mathbb{1}\left\{r_{\theta^\star}(x,y) + \lambda(x) > 0\right\}\right) > 0$). If $\mathbb{E}[r_{\theta^\star}(x,y)] < 0$, it is saying the overpayments and selection rule need to be much more associated to overcome the average underpayment. In general, somehow the overpayments and selection indicator need to move together (and the strength of how much they need to move together depends on overpayments on the population.

Example: Rademacher variables. For simplicity, we will consider scalar x. Suppose $x \in \{\pm 1\}$ with $\Pr(x = -1) = \Pr(x = +1) = 1/2$. Then $y = x\beta + \varepsilon$ where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Let $\lambda(x) = x\alpha$. We then have

$$\begin{aligned}
r_{\theta^\star}(x,y) &= x(\theta^\star - \beta) - \varepsilon \\
r_{\theta^\star}(x,y) + \lambda(x) &= x\left(\theta^\star - \beta + \alpha\right) - \varepsilon \\
\pi_{\theta^\star}(x,y) &= \mathbb{1}\left\{\varepsilon < x\left(\theta^\star - \beta + \alpha\right)\right\}\,.
\end{aligned}$$

Using the optimality of $\theta^\star$ on $\overline{\pi^\star}$:

$$\begin{aligned}
&\mathbb{E}\left[x(y - x\theta^\star) \mid \varepsilon \geq x(\theta^\star - \beta + \alpha)\right] = 0 \\
\Rightarrow\ &\mathbb{E}\left[x(x\beta + \varepsilon - x\theta^\star) \mid \varepsilon \geq x(\theta^\star - \beta + \alpha)\right] = 0 \\
\Rightarrow\ &\mathbb{E}\left[x^2(\beta - \theta^\star) + x\varepsilon \mid \varepsilon \geq x(\theta^\star - \beta + \alpha)\right] = 0 \\
\Rightarrow\ &\mathbb{E}\left[(\beta - \theta^\star) + x\varepsilon \mid \varepsilon \geq x(\theta^\star - \beta + \alpha)\right] = 0 \qquad \text{(using } x^2 = 1\text{)}
\end{aligned}$$

Then

$$\mathbb{E}[x\varepsilon \mid \overline{\pi^\star}] = \mathbb{E}\left[\beta - \theta^\star \mid \overline{\pi^\star}\right]\,. \quad (17)$$

But

$$\mathbb{E}\left[x\varepsilon \mid \overline{\pi^\star}\right] = \mathbb{E}_x\left[\frac{\mathbb{E}\left[x\varepsilon\mathbb{1}\left\{\varepsilon \geq x(\theta^\star - \beta + \alpha)\right\} \mid x\right]}{\Pr\left(\varepsilon \geq x(\theta^\star - \beta + \alpha) \mid x\right)}\right]\,. \quad (18)$$

Conditioning on x,

$$\begin{aligned}
\mathbb{E}\left[\varepsilon\mathbb{1}\left\{\varepsilon \geq x(\theta^\star - \beta + \alpha)\right\}\right] &= \sigma\phi\left(\frac{x(\theta^\star - \beta + \alpha)}{\sigma}\right) \\
\mathbb{E}\left[x\varepsilon\mathbb{1}\left\{\varepsilon \geq x(\theta^\star - \beta + \alpha)\right\}\right] &= \frac{1}{2}\left(\sigma\phi\left(\frac{x(\theta^\star - \beta + \alpha)}{\sigma}\right) - \sigma\phi\left(\frac{-x(\theta^\star - \beta + \alpha)}{\sigma}\right)\right) = 0\,. \\
&\qquad \text{(using } \phi(z) = \phi(-z)\text{)}
\end{aligned}$$

This implies that $\mathbb{E}[x\varepsilon \mid \overline{\pi^\star}] = 0$ by Equation (18). Thus $\theta^\star = \beta$ by Equation (17). Then we actually have that

$$\begin{aligned}
r_{\theta^\star}(x,y) &= -\varepsilon \\
\Rightarrow \mathbb{E}\left[r_{\theta^\star}(x,y)\cdot\pi^\star(x,y) \mid x\right] &= \mathbb{E}\left[-\varepsilon\cdot\mathbb{1}\left\{\varepsilon < x\alpha\right\} \mid x\right] = \sigma\phi\left(\alpha x/\sigma\right) = \sigma\phi(\alpha/\sigma) \\
&\qquad \text{(using } x = \pm 1 \text{ and } \phi(z) = \phi(-z)\text{)} \\
\Rightarrow E\left[r_{\theta^\star}(x,y)\cdot\pi^\star(x,y)\right] &= \sigma\phi(\alpha/\sigma) > 0\,.
\end{aligned}$$

Also, $\Pr(\overline{\pi^\star}) = \frac{1}{2}\Pr\left(\varepsilon \geq x\alpha\right) + \frac{1}{2}\Pr\left(\varepsilon \geq -x\alpha\right) = \frac{1}{2} > 0$. Thus $E\left[r_{\theta^\star}(x,y) \mid \pi^\star(x,y) = 1\right] > 0$.

D.3 Convergence Results

Bounded $\lambda(x)$. In this section, we discuss what kind of convergence results can be possible under additional assumptions. We point out that while in the main body we made the assumption that $\lambda(x) = \alpha^\top x + \alpha_0$, assuming a bound on $\lambda(x)$ conflicts with the linear definition. However, a closer look at Lemma C.1 and Theorems 1 and 2 shows that many of the qualitative conclusions hold regardless of the form of $\lambda(x)$. The OLS optimality on $\pi^\star$ leads to no government overpayment (Phenomenon 8) under the reformed policy, the organization's margin is always positive regardless of the equilibrium due to Phenomenon 3, and we can always decompose the organization's margin as the government overpayment and the efficient from $\lambda(x)$ (Phenomenon 4 and Phenomenon 9). One may also be able to clip $\lambda(x) = \max\{\Lambda, \alpha^\top x + \alpha_0\}$ which may approximate the affine case depending on the magnitude of α, α_0, but we do not investigate this.

Proposition D.2 (Dynamics updates under Equation (4) are a contraction under bounded $\lambda(x)$)**.** *Let $|\lambda(x)| \le \Lambda$ for all x. Define*

$$L_w(\theta) \coloneqq \mathbb{E}\left[\left(y - x^\top\theta\right)^2 \mid w, x^\top w \le y - \lambda(x)\right], \qquad T(w) \coloneqq \arg\min_\theta L_w(\theta)\,.$$

Thus $\theta^{t+1} = T(\theta^t)$ for every iteration t of the dynamics described by Equation (4). Let $w^\star$ be a fixed point so $T(w^\star) = w^\star$. If

$$\mathbb{E}_x\left[\Phi\left(-\frac{x^\top(w^\star - \beta) + \lambda(x)}{\sigma}\right) xx^\top\right] \succeq \mu I\,, \quad \textit{and} \quad \Lambda < \frac{\mu\sigma\sqrt{2\pi}}{\|\mathbb{E}_x[xx^\top]\|_{2\to 2}}\,,$$

then

$$\|\nabla_w T(w^\star)\|_{2\to 2} < 1\,,$$

i.e., the map $T(w)$ is locally contracting near $w^\star$.

Remark 7. Looking at the conditions we need, the covariance bound is saying that there is enough variability in covariates on the complement of the selected set at equilibrium. Another interpretation is that selection does not cause the complement set to collapse. This may be reasonable to hope for if we want a finite, non-boundary equilibrium. Otherwise, if $\lambda(x) \ge 0$ for all x, for instance, we do converge to a boundary equilibrium where $\pi^\star(x, y) \equiv 1$ and $\theta^\star$ is undefined (see Section D.1). The condition on the magnitude of $\lambda(x)$ can roughly be compared to σ, and it is reasonable to want the gap between y and y' to be less than the uncertainty in the y measurements themselves (up to some factors). Without any bound on $\lambda(x)$ for instance, note that for any x, $\Pr\left(x^\top\theta > y - \lambda(x) \mid x\right) = \Phi\left(1/\sigma \cdot (x^\top(\theta - \beta) + \lambda(x))\right)$. As $\lambda(x) \to -\infty$, this probability goes to 0, i.e., if all $\lambda(x)$ are largely negative then it is highly unlikely to see any points in the selected set in the first place.

Proof. **Step 1: Solving for $T(w)$.** Note that

$$\begin{aligned} T(w) &= \arg\min_\theta \mathbb{E}\left[\left(y - x^\top\theta\right)^2 \mid w, x^\top w \le y - \lambda(x)\right] \\ &= \arg\min_\theta \frac{\mathbb{E}\left[\mathbb{1}\left\{x^\top w \le y - \lambda(x)\right\} \cdot \left(y - x^\top\theta\right)^2 \mid w\right]}{\Pr\left(x^\top w \le y - \lambda(x)\right)} \\ &= \arg\min_\theta \mathbb{E}\left[\mathbb{1}\left\{x^\top w \le y - \lambda(x)\right\} \cdot \left(y - x^\top\theta\right)^2 \mid w\right]. \\ &\qquad\qquad \text{(as } \Pr\left(x^\top w \le y - \lambda(x)\right) \text{ does not depend on } \theta) \end{aligned}$$

By the first-order optimality condition for OLS, we have

$$\begin{aligned} \nabla_\theta L_w(\theta) &= \mathbb{E}\left[\mathbb{1}\left\{x^\top w \le y - \lambda(x)\right\} \cdot \left(-2x\left(y - x^\top\theta\right)\right)\right] \\ &= \mathbb{E}\left[-2xy \cdot \mathbb{1}\left\{\varepsilon \ge x^\top(w - \beta) + \lambda(x)\right\}\right] + \mathbb{E}\left[2xx^\top\theta \cdot \mathbb{1}\left\{\varepsilon \ge x^\top(w - \beta) + \lambda(x)\right\}\right]. \\ &\qquad\qquad \text{(using } y = x^\top\beta + \varepsilon) \end{aligned}$$

Solving for θ when $\nabla_\theta L_w(\theta) = 0$:

$$\Rightarrow \mathbb{E}\left[xx^\top\theta \cdot \mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right] = \mathbb{E}\left[xy \cdot \mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right]$$

$$\Rightarrow \mathbb{E}\left[\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}xx^\top\right]\theta = \mathbb{E}\left[xy \cdot \mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right]$$

$$\Rightarrow \theta = \left(\mathbb{E}\left[\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\} \cdot xx^\top\right]\right)^{-1}\mathbb{E}\left[\mathbb{1}\left\{\varepsilon > x^\top(w-\beta)+\lambda(x)\right\} \cdot xy\right]$$

$$\Rightarrow \theta = \left(\mathbb{E}_x\left[\underbrace{\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot xx^\top}_{\Pr(\varepsilon \geq x^\top(w-\beta)+\lambda(x))}\right]\right)^{-1}\mathbb{E}_x\left[\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot xy\right]$$

$$\Rightarrow \theta = \mathbb{E}_x\left[\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) \cdot xx^\top\right]^{-1}\mathbb{E}_x\left[\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot xy\right]$$

Specifically for the second term:

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot xy\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot x\left(x^\top\beta+\varepsilon\right)\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_\varepsilon\left(\mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right) \cdot xx^\top\beta\right] + \mathbb{E}_x\left[x \cdot \mathbb{E}_\varepsilon\left[\varepsilon \cdot \mathbb{1}\left\{\varepsilon \geq x^\top(w-\beta)+\lambda(x)\right\}\right]\right] \\
&= \mathbb{E}_x\left[\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) \cdot xx^\top\beta\right] + \mathbb{E}_x\left[x \cdot \left(\sigma\phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right)\right].
\end{aligned}$$

Plugging this back into our expression for θ which is what $T(w)$ returns:

$$T(w) = \beta + \mathbb{E}_x\left[\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) \cdot xx^\top\right]^{-1}\mathbb{E}_x\left[x \cdot \left(\sigma\phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right)\right].$$

Step 2: Jacobian has bounded operator norm. Defining:

$$A(w) = \mathbb{E}_x\left[\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) \cdot xx^\top\right] \qquad b(w) = \mathbb{E}_x\left[x \cdot \left(\sigma\phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right)\right],$$

then we can rewrite

$$T(w) = \beta + A(w)^{-1}b(w),$$

and the Jacobian of $T(w)$ is described by

$$\begin{aligned}
\frac{\partial T(w)}{\partial w_i} &= \left(\frac{\partial A(w)^{-1}}{\partial w_i}\right)b(w) + A(w)^{-1}\left(\frac{\partial b(w)}{\partial w_i}\right) \\
&= \left(-A(w)^{-1}\mathbb{E}_x\left[\frac{\partial}{\partial w_i}\left(\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right) \cdot xx^\top\right]A(w)^{-1}\right)b(w) \\
&\qquad + A(w)^{-1}\left(\mathbb{E}_x\left[x \cdot \frac{\partial}{\partial w_i}\left(\sigma\phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right)\right]\right),
\end{aligned}$$

for each component $i \in [d]$. Further,

$$\frac{\partial}{\partial w_i}\Phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) = -\frac{x_i}{\sigma}\phi\left(-\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right), \qquad \text{(using } \Phi'(z) = \phi(z)\text{)}$$

and

$$\frac{\partial}{\partial w_i}\left(\sigma\phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right)\right) = -x_i \cdot \left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right) \cdot \phi\left(\frac{x^\top(w-\beta)+\lambda(x)}{\sigma}\right).$$
(using $\phi'(z) = -z\phi(z)$)

Let $r_w(x) = \frac{x^\top(w-\beta)+\lambda(x)}{\sigma}$. Then we have

$$\begin{aligned}
\frac{\partial T(w)}{\partial w_i} &= -A(w)^{-1}\left(-\frac{1}{\sigma}\mathbb{E}_x\left[\phi\left(-r_w(x)\right)x_i xx^\top\right]\underbrace{A(w)^{-1}b(w)}_{T(w)-\beta} - \mathbb{E}_x\left[x_i x \cdot \left(r_w(x)\right)\phi\left(r_w(x)\right)\right]\right) \\
&= A(w)^{-1}\mathbb{E}_x\left[\frac{x_i}{\sigma}\phi(r_w(x))\left(xx^\top(T(w)-\beta) - \sigma x r_w(x)\right)\right] \\
\Rightarrow \frac{\partial T(w)}{\partial w} &= A(w)^{-1}\mathbb{E}_x\left[\frac{1}{\sigma}\phi(r_w(x)) \cdot x\left(x^\top(T(w)-\beta) - \sigma r_w(x)\right)x^\top\right].
\end{aligned}$$

Let $w^\star$ be a fixed point so that $T(w^\star) = w^\star$. Then $\frac{1}{\sigma}x^\top (T(w^\star) - \beta) = \frac{1}{\sigma}x^\top(w^\star - \beta) = r_{w^\star}(x) - \lambda(x)/\sigma$, so

$$\left.\frac{\partial T(w)}{\partial w}\right|_{w=w^\star} = -A(w^\star)^{-1}\mathbb{E}_x\left[\phi(r_{w^\star}(x)) \cdot \frac{\lambda(x)}{\sigma} \cdot xx^\top\right].$$

Then we have that

$$\begin{aligned}
\|\nabla T(w^\star)\|_{2\to 2} &\le \left\|A(w^\star)^{-1}\right\|_{2\to 2} \left\|\mathbb{E}_x\left[\frac{1}{\sigma\sqrt{2\pi}}\Lambda \cdot xx^\top\right]\right\|_{2\to 2} \\
&\qquad \text{(using } \phi(\cdot) \le 1/\sqrt{2\pi} \text{ and } \lambda(x) \le \Lambda \text{ for all } x\text{)} \\
&\le \frac{\Lambda}{\mu\sigma\sqrt{2\pi}} \left\|\mathbb{E}_x[xx^\top]\right\|_{2\to 2} \qquad \text{(using } \left\|A(w^\star)^{-1}\right\| \le 1/\mu\text{)}
\end{aligned}$$

This is bounded by 1 provided $\Lambda < \frac{\mu\sigma\sqrt{2\pi}}{\|\mathbb{E}_x xx^\top\|}$. □

Once we have a locally contracting map, it is straightforward to say that convergence happens quickly once we enter a neighborhood of the fixed points.

Corollary D.3 (Convergence of dynamics in Equation (4))**.** *Suppose $\nabla T(w)$ is continuous in w. Let $\|\nabla T(w^\star)\| = \rho^\star < 1$. By continuity, there exists $r > 0$, $\rho^\star \le \rho < 1$ such that for all $w \in \mathcal{B}(w^\star, r)$, $\|T(w)\|_{2\to 2} \le \rho < 1$. Then it holds that if w^0 is initialized in $\mathcal{B}(w^\star, r)$,*

$$\left\|w^t - w^\star\right\|_2 \le \rho^t \left\|w^0 - w^\star\right\|_2 \le \rho^t r\,.$$

Remark 8. The continuity assumption may be satisfied, e.g., if $\mathbb{E}\|x\|_2^m < \infty$ for a large enough (finite) m so that we can bound the terms in appearing in $\nabla T(w)$ to apply dominated convergence. This applies to Gaussian x. We can prove a more global convergence result if we allow ourselves stronger assumptions. In particular, if we assume the same conditions on the selection-weighted covariance matrix and $\Pr\left(-(x^\top(w-\beta) + \lambda(x))/\sigma\right) > 0$ for every w uniformly in a (subset of the) parameter space (not just at $w^\star$), we can bound $\|\nabla T(w)\|_{2\to 2}$ uniformly over this space.

Remark 9 (Similar convergence results under reformed policy)**.** We can prove similar convergence results under the reformed policy, where changes mainly occur due to the reversed inequalities (upper vs. lower Gaussian tails).

D.4 Nonparametric Regression

Instead of a linear regression model $f_\theta(x) = x^\top\theta$ consider a non-parametric model $f : \mathcal{X} \to \mathbb{R}$. The best responses at a given round t for the max player and the min player are, respectively

$$\pi^t(x, y') = \mathbb{1}\left\{y' < f^t(x)\right\} \quad \text{and} \quad f^{t+1}(x) = \frac{\mathbb{E}\left[y \cdot (1 - \pi^t(x, y')) \mid x\right]}{\mathbb{E}\left[1 - \pi^t(x, y') \mid x\right]} = \mathbb{E}\left[y \mid y' \ge f^t(x), x\right].$$

But on the selected population, we have

$$\begin{aligned}
\mathbb{E}\left[f^{t+1}(x) - y \mid y' < f^t(x), x\right] &= f^{t+1}(x) - \mathbb{E}[y \mid y' < f^t(x), x] \\
&= \mathbb{E}\left[y \mid y' \ge f^t(x) \mid x\right] - \mathbb{E}[y \mid y' < f^t(x), x] \\
&\qquad \text{(definition of } f^{t+1}\text{)}
\end{aligned}$$

The overpayment is positive exactly under the condition:

$$\mathbb{E}\left[y \mid y' \ge f^t(x), x\right] > \mathbb{E}\left[y \mid y' < f^t(x), x\right]. \tag{19}$$

D.4.1 Sufficient Conditions

Intuitively, we can see just from Equation (19) that the quantities y and y' must be positively correlated so that a thresholding of one causes the expected value of the other to be affected.

Positive Covariance Between y and selection indicator. Define the following:

$$\eta \coloneqq f^t(x)\,, \quad S_\eta \coloneqq \mathbb{1}\{y' \geq \eta\}\,, \quad \text{and} \quad p \coloneqq \Pr(S_\eta = 1 \mid x)\,.$$

Assume that $0 < p < 1$ so both conditional means are well-defined in Equation (19). By the law of total expectation,

$$\mathbb{E}[y \mid x] = p \cdot \mathbb{E}[y \mid S_\eta = 1, x] + (1-p) \cdot \mathbb{E}[y \mid S_\eta = 0, x]\,.$$

The covariance between y and S_η (selection indicator) is

$$\begin{aligned}
\mathrm{Cov}(y, S_\eta) &= \mathbb{E}[yS_\eta \mid x] - \mathbb{E}[y \mid x] \cdot \mathbb{E}[S_\eta \mid x] \\
&= \Pr(S_\eta = 1 \mid x)\,\mathbb{E}[y \mid S_\eta = 1, x] - \Big(p \cdot \mathbb{E}[y \mid S_\eta = 1, x] + (1-p) \cdot \mathbb{E}[y \mid S_\eta = 0, x]\Big) \cdot p \\
&= p(1-p) \cdot \mathbb{E}[y \mid S_\eta = 1, x] - (1-p)p \cdot \mathbb{E}[y \mid S_\eta = 0, x] \\
&= p(1-p) \cdot \Big(\mathbb{E}[y \mid S_\eta = 1, x] - \mathbb{E}[y \mid S_\eta = 0, x]\Big)\,.
\end{aligned}$$

Since $0 < p < 1$, $\mathbb{E}[y \mid S_\eta = 1, x] > \mathbb{E}[y \mid S_\eta = 0, x] \iff \mathrm{Cov}(y, S_\eta) > 0$. Note that $\mathbb{E}[y \mid S_\eta = 1, x] > \mathbb{E}[y \mid S_\eta = 0, x]$ is equivalent to the overpayment condition in Equation (19). This type of condition is similar to what we saw in Equation (16).

Monotonicity. The most general sufficient condition is an abstract monotonicity statement. If we have the property that

$$\mathbb{E}[y \mid y' = \tau, x] \text{ is increasing (non-decreasing) in } \tau\,,$$

then clearly, the condition in Equation (19) holds regardless of $f^t(x)$ for every x. However, this is somewhat abstract and unhelpful unless we have other more interpretable conditions which can imply this.

Conditional Bivariate Normality. In general, positive covariance between y and y' (rather than y and the selection indicator) is not sufficient. However, if y and y' are jointly Gaussian, we say the following. Suppose

$$(y, y') \mid x \sim \mathcal{N}(\mu, \Sigma)\,.$$

For bivariate normal, the conditional expectation of y given y' is given by:

$$\mathbb{E}[y \mid y' = \tau, x] = \mu_y(x) + \frac{\mathrm{Cov}(y, y' \mid x)}{\mathrm{Var}(y' \mid x)}\,(\tau - \mu_{y'}(x))\,. \tag{20}$$

With $\mathrm{Cov}(y, y' \mid x) > 0$, the coefficient in front of the term involving τ is positive and thus $\mathbb{E}[y \mid y' = \tau, x]$ is increasing in τ.

D.4.2 Both y and y' depend on x linearly

Suppose we have the following:

$$y = w^\top x + \varepsilon, \qquad y' = v^\top x + \xi, \qquad \text{where} \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)\,, \xi \sim \mathcal{N}(0, \sigma'^2)\,, \mathrm{Cov}(\varepsilon, \xi) = \rho\,.$$

Then indeed, y and y' are jointly Gaussian, conditioned on x. Note that if y and y' are conditionally independent, Equation (20) says $\mathbb{E}[y \mid y', x] = \mu_y(x)$. This is expected since if y and y' are conditionally independent, the change in one variable should have no effect on the other. Thus we assume $\rho > 0$.

Then Equation (20) becomes

$$\mathbb{E}[y \mid y' = \tau, x] = w^\top x + \frac{\rho}{\sigma'} \cdot (\tau - v^\top x) = \left(w + \frac{\rho}{\sigma'} v\right)^\top x + \frac{\rho}{\sigma'} \cdot \tau\,.$$

D.4.3 Explicit dependence as $y' = y - \lambda(x)$

If $y' = y - \lambda(x)$, this is equivalent to

$$\mathbb{E}\left[y \mid y \geq f^t(x) + \lambda(x), x\right] > \mathbb{E}\left[y \mid y < f^t(x) + \lambda(x), x\right]\,.$$

This should always be true, given that the conditional distribution of y given x is non-degenerate.

Existence of fixed point. At equilibrium, the condition for $f^\star$ is

$$f^\star(x) = \mathbb{E}\left[y \mid y \geq f^\star(x) + \lambda(x), x\right] .$$

Unless $\lambda(x) \leq 0$, this is hard to satisfy point-wise. Thus there may not exist a (finite) fixed point for this setting for general $\lambda(x)$. Otherwise, when $\lambda(x) \geq 0$, this is tending to infinity.

E Simulation Details

Figure 2. We generate $n = 50$ observations with $x \sim \mathcal{N}(0, 1)$ and $y = 0.5 + \beta x + \varepsilon$, where $\beta \in \{-2, 2\}$ and $\varepsilon \sim \mathcal{N}(0, 49)$ is independent Gaussian noise. We set $\lambda(x) = 0$, so that $y' = y$, and initialize the payment rule using OLS on all observations. We run the status quo dynamics for up to 10 iterations, refitting on the unselected observations at each step.

Figures 3 and 4. We generate $n = 1000$ observations with $d = 3$ features drawn independently from $\mathcal{N}(0, 1)$. Outcomes are generated according to $y = \beta_0 + x^\top \beta + \varepsilon$, where $\beta_0 = -1.5$, $\beta = (0.5, 1.2, 3.0)$, and $\varepsilon \sim \mathcal{N}(0, 25)$. We set $\lambda(x) = \alpha_0 + x^\top \alpha$, with $\alpha = (1, -2, 0.75)$, and vary the intercept parameter $\alpha_0 \in \{2, 1, 0, -1, -2\}$ to study different baseline efficiency advantages. Both policies are initialized using OLS on the full population and run for up to 30 iterations.

F Experimental Details and Additional Experiments

F.1 Data Description

We use the `upcoding`[2] package to simulate a realistic distribution of co-occurring HCCs (diagnosis groups) in the Medicare population. The package is calibrated using self-reported health conditions from Medicare-eligible participants aged 65 years and older in the National Institutes of Health's All of Us study, which are less affected by coding incentives than claims-based data [61].

To obtain risk scores for the beneficiaries, we use the officially published CMS-HCC model coefficients[3]. To map these risk scores to realistic monetary values, we use an annual benchmark of 13261.2, derived from the 2024 projected national per capita Fee-For-Service Medicare spending (USPCC) rate of \$1,105.10 per month[4]. Specifically, we generate annual costs as

$$y = 13{,}261.20 \cdot \max\{0.01, \theta_0 + x^\top \theta^*\} \cdot \exp(0.5\varepsilon - 0.125), \qquad \varepsilon \sim \mathcal{N}(0, 1),$$

where θ^* contains the CMS-HCC coefficients and $\theta_0 = 0.35$ approximates the base demographic relative factor for a typical 70-year-old beneficiary. The multiplicative log-normal term captures the heavy right tail of healthcare spending [62], while ensuring that expected costs remain proportional to the risk score. The lower bound of 0.01 ensures positive risk scores but does not affect any observations in our simulated data.

In Figure 7, we plot the distribution of annual medical costs (in \$) in our simulated data and the distribution of HCCs (diagnosis groups) per person. In Figure 8, we plot the top 20 diagnosis groups based on their prevalence in the population.

F.2 Additional Experimental Details

We use the coefficients θ^* published by CMS for their V28 CMS-HCC model, together with the fixed demographic factor θ_0 described above.

We construct an affine function $\lambda(x) = \alpha_0 + x^\top \alpha$, where both the intercept and feature effects are scaled to dollars using the Medicare spending benchmark. Since $\lambda(x) = y - y'$, a positive value of $\lambda(x)$ means that the insurer's realized service cost is lower than the service cost under the existing FFS program. In our experiments, $\alpha_0 = 0.01 \times 13{,}261.20$ represents a baseline efficiency advantage for the insurer. The vector α then introduces heterogeneity across clinical conditions: for a randomly

[2] https://github.com/StanfordHPDS/upcoding

[3] https://www.cms.gov/files/document/2024-announcement-pdf.pdf

[4] https://www.commonwealthfund.org/publications/explainer/2024/mar/how-government-updates-payment-rates-medicare-advantage-plans

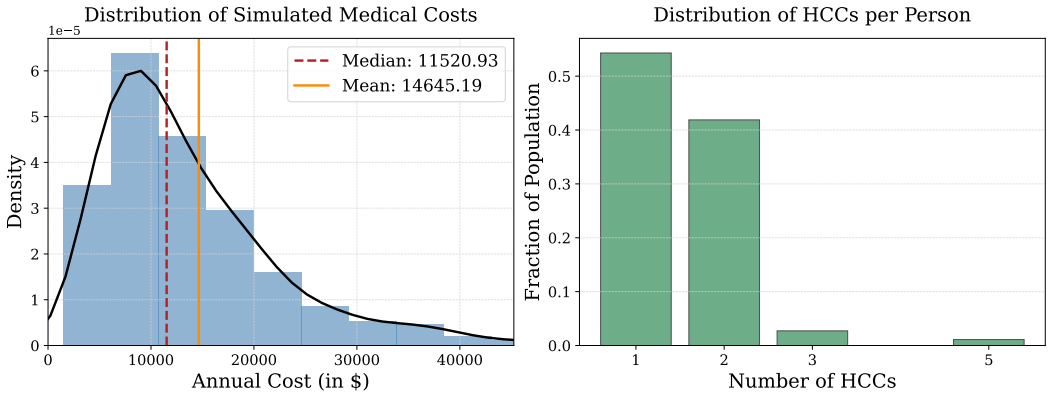

Figure 7: Distribution of simulated annual medical costs (in $) (left) and distribution of HCCs (diagnosis groups) per beneficiary (right).

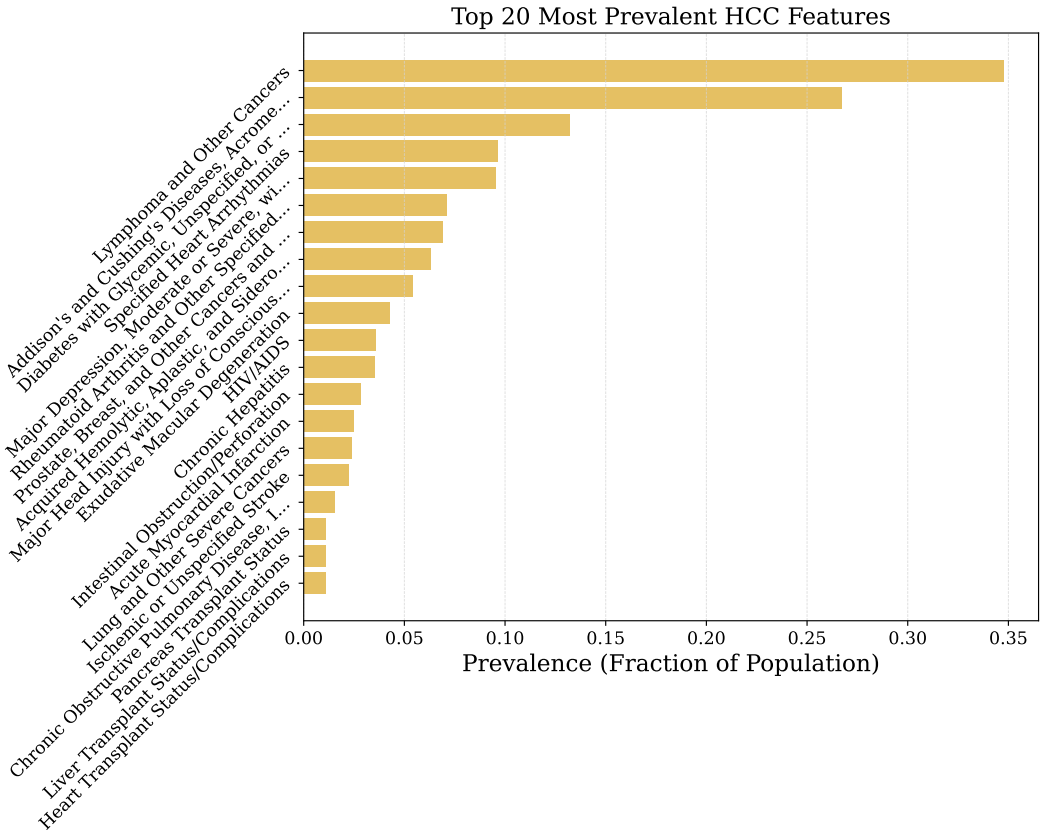

Figure 8: **Top 20 most prevalent HCC features (diagnosis groups) in the simulated beneficiary population**. Prevalence is measured by the proportion of beneficiaries with each HCC.

chosen subset of HCCs, $\alpha_j > 0$ represents more favorable conditions whose presence increases insurer cost advantages, while for another subset, $\alpha_j < 0$ represents conditions associated with smaller or negative cost advantages, such as administrative burden or less efficient service delivery. The remaining conditions have $\alpha_j = 0$, so they do not shift costs relative to baseline. Specifically, 60% of features are assigned positive values sampled uniformly from $[0.03, 0.12]$, 20% are assigned negative values from $[-0.05, -0.01]$, and the remaining 20% are set to zero, with all coefficients scaled by the annual spending benchmark

For both policies, we initialize the payment rule using OLS on the full population. At each iteration, the insurer selects beneficiaries whose predicted payments exceed its service costs. The decision maker then refits the payment rule on the unselected population under the status quo and on the selected population under the reform, using the simulated counterfactual costs y. We run the dynamics for up to 100 iterations, topping early if the payment coefficients converge and the selected set remains unchanged.

F.3 Additional Results

In Figure 9, we plot average payments and true costs (in $) among selected beneficiaries under the status quo and reformed policies. Even though the data do not match the distributional assumptions used in our theory, we observe similar qualitative behavior: overpayment becomes substantial and continues to fluctuate under the status quo, while it converges to zero under the reformed policy. The reform reaches a fixed point after 22 iterations, whereas the status quo does not converge within 100 iterations.

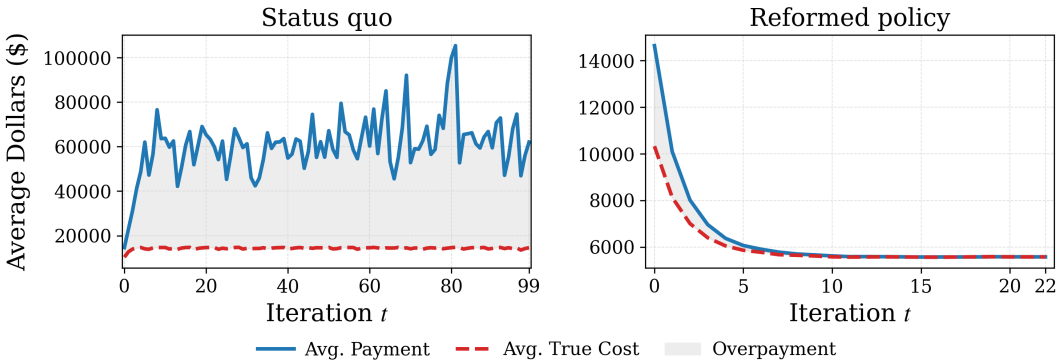

Figure 9: **Average payment vs average true cost (in $) under the status quo (left) and reformed policies (right).** Solid and dashed lines represent average payments and true costs among selected beneficiaries, respectively, with the shaded region indicating overpayment.

NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes], [No], or [N/A].
- [N/A] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for [N/A]).

The checklist answers are an integral part of your paper submission. They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will also be asked to include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While [Yes] is generally preferable to [No], it is perfectly acceptable to answer [No] provided a proper justification is given (e.g., error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering [No] or [N/A] is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS Paper Checklist"**,
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers**.

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: We claim that (1) we provide a theoretical framework for understanding empirical observations in risk-based allocation. This is explained in Section 2. We then analyze status quo policies in Section 3.2 and propose a policy reform in Section 4. We provide proofs for all our claims, which are often deferred to the Appendix.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [Yes]

Justification: We clearly list that some of our theoretical results do not explain simulation results and limitations of fixed point analyses in the discussion and future work section Section 6.

Guidelines:

- The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: Each theorem statement and claim lists the assumptions needed at the beginning and proofs are included in the Appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

Justification: We provide the simulation and experimental details in the main paper (Section 5) and appendix (Appendix F).

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: We provide these in the supplementary material.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.

- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: We explain in Section 5 and Section F.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: In our experiments, it is not important that any particular error metric is significant compared to a baseline. We simulate dynamics to draw conclusions about broader patterns.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [No]

Justification: The simulations are lightweight and can be done on, e.g., the free tier of a Google Colab notebook.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: We conform, in every respect, with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Throughout the manuscript, our focus is on societal impacts of existing policies in risk-based allocation. We make efforts to comment on such broader impacts throughout, e.g., Section 1,Section 6.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: We do not release such data or models.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We use the upcoding package (https://github.com/StanfordHPDS/upcoding) to generate data for experiments and we cite the authors in the manuscript as well as include the URL.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: We do not introduce any significant new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.

- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: Our paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: Our paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.